Biochemical pathways

What is a “pathway”

Wikipedia (October 14th 2013): “In biochemistry, metabolic pathways are series of chemical reactions occurring within a cell. In each pathway, a principal chemical is modified by a series of chemical reactions. Enzymes catalyze these reactions [...]”

Different types: Signalling pathways, metabolic networks, gene regulatory networks ...

Many “pathway” databases:

Biocarta, Bio/MetaCyc, Ingenuity IPA, KEGG Pathway, Panther pathways, Reactome, STKE, Wikipathways etc.

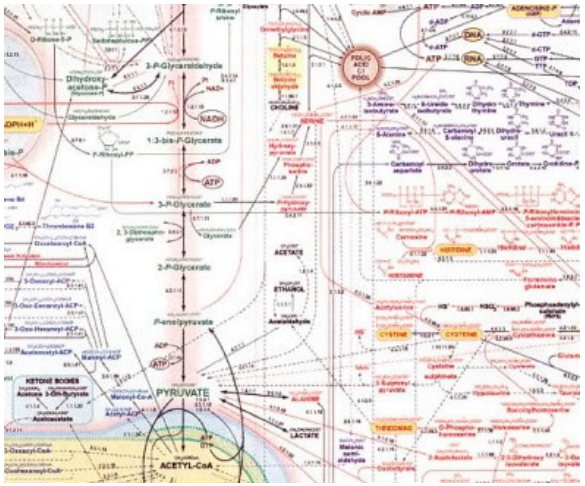
➜ Detailed representation of reality based on observation

OR

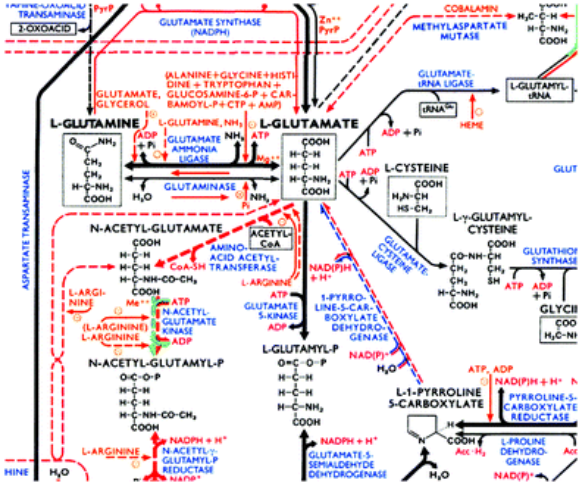
➜ Abstract representation of reality based on needs

Biochemical pathways are old …

- Gortner, R.A. *Outlines of Biochemistry* (Wiley, New York, 1949)
- Nicholson (1970)

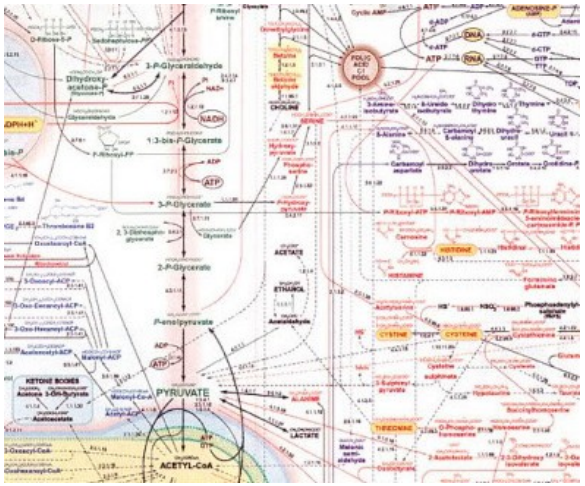


Michal (1984)

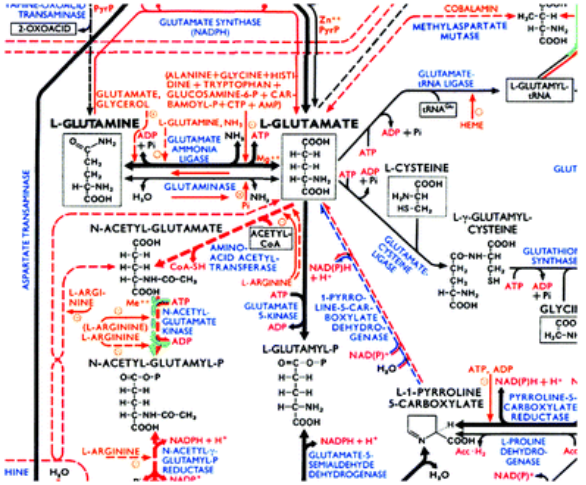


Biochemical pathways are old …

- Gortner, R.A. *Outlines of Biochemistry* (Wiley, New York, 1949)
- Nicholson (1970)



Michal (1984)

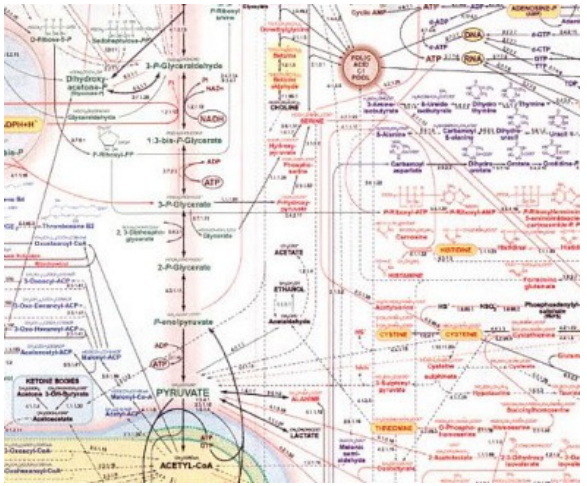


"Hand drawing" on paper

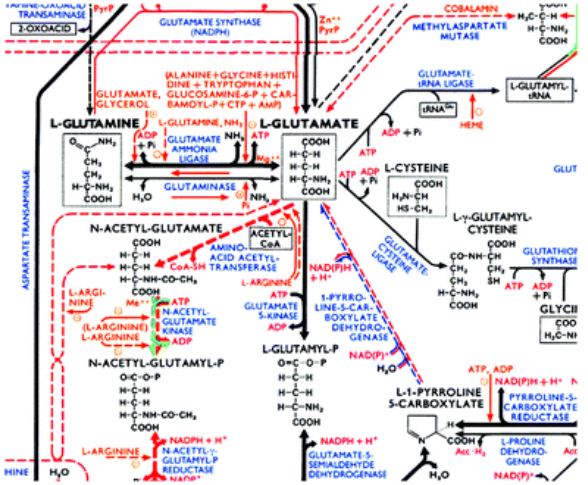
⟶ no software-based browsing, processing and analysis

Biochemical pathways are old … or not so much

- Gortner, R.A. *Outlines of Biochemistry* (Wiley, New York, 1949)
- Nicholson (1970)

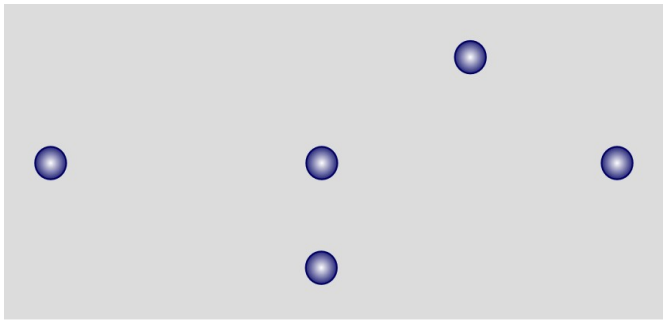


Michal (1984)

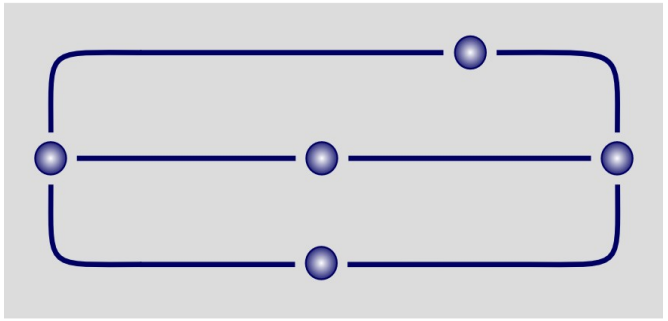


- 1990s: high-throughput data generation → Large amount of knowledge
  increase in computing power → automatic reconstruction, browsing and analysis
- Databases: KEGG (1995), EcoCyc (1994), Reactome (2000)
- Formats: BioPAX (2000), SBML (2000), PSI-MI (2002)

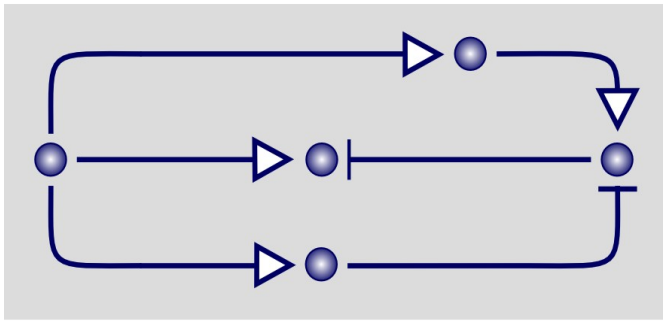
Layered approach for designing pathways



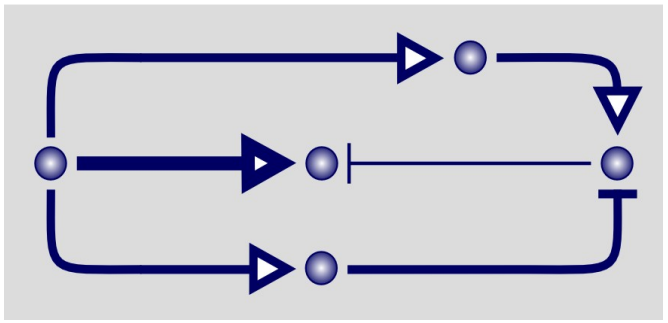
List of participants



Interactions between participants



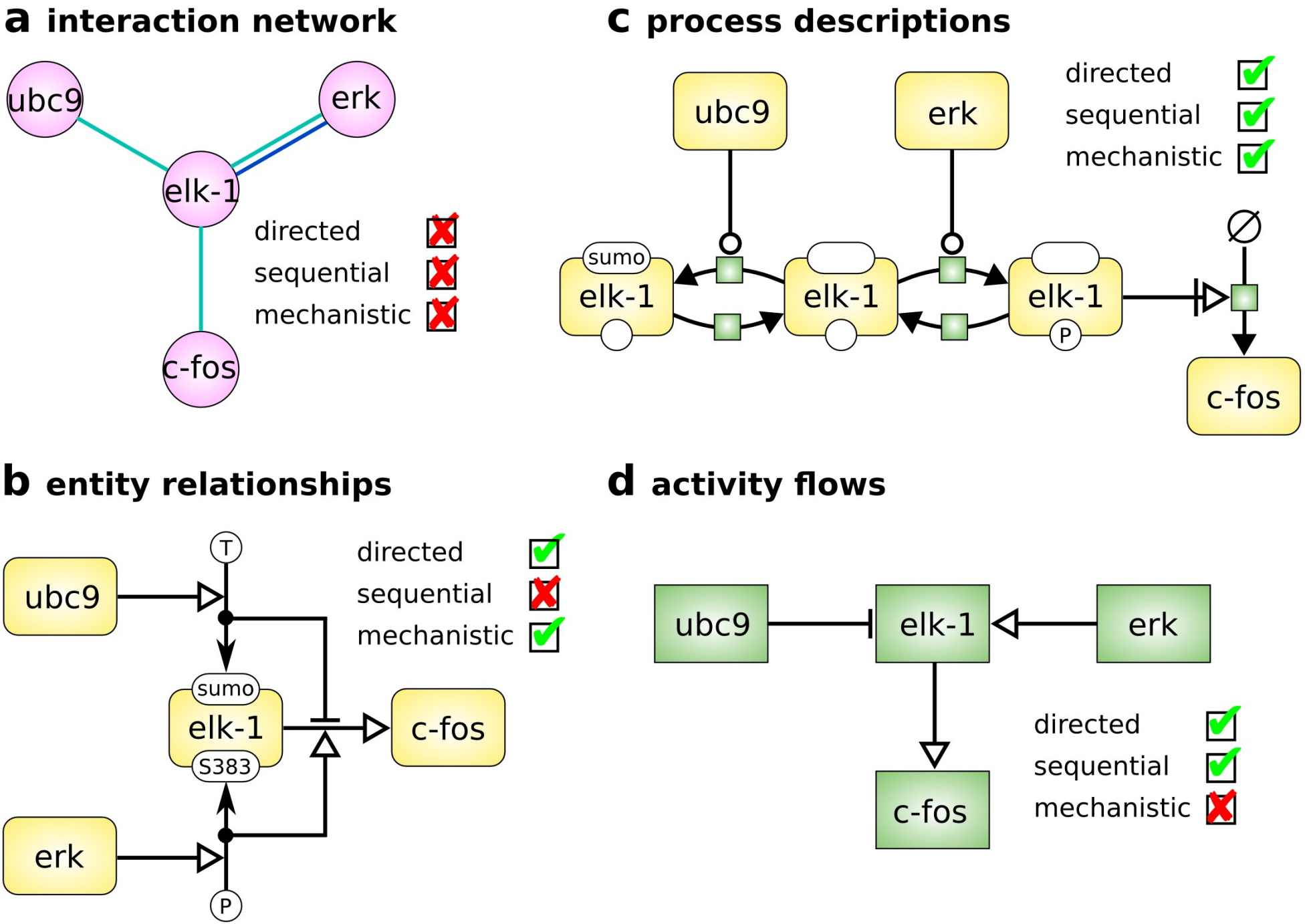
Influences of participants onto others

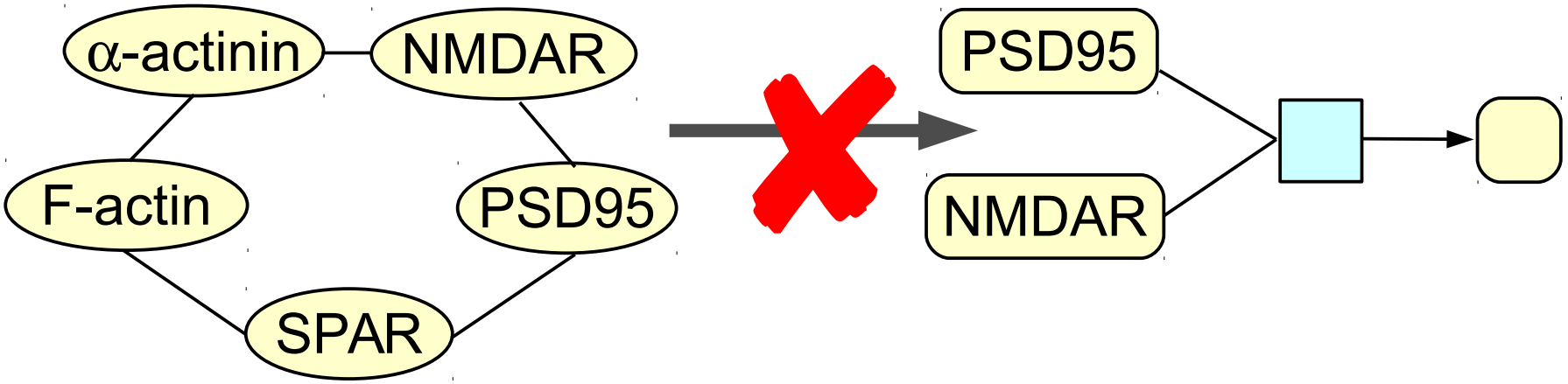


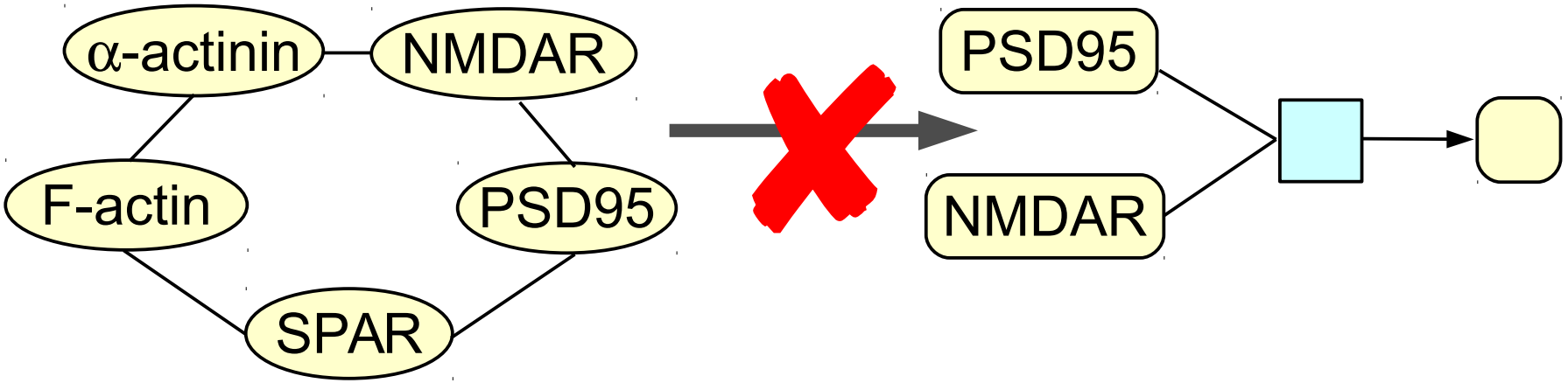
Quantitative relationships

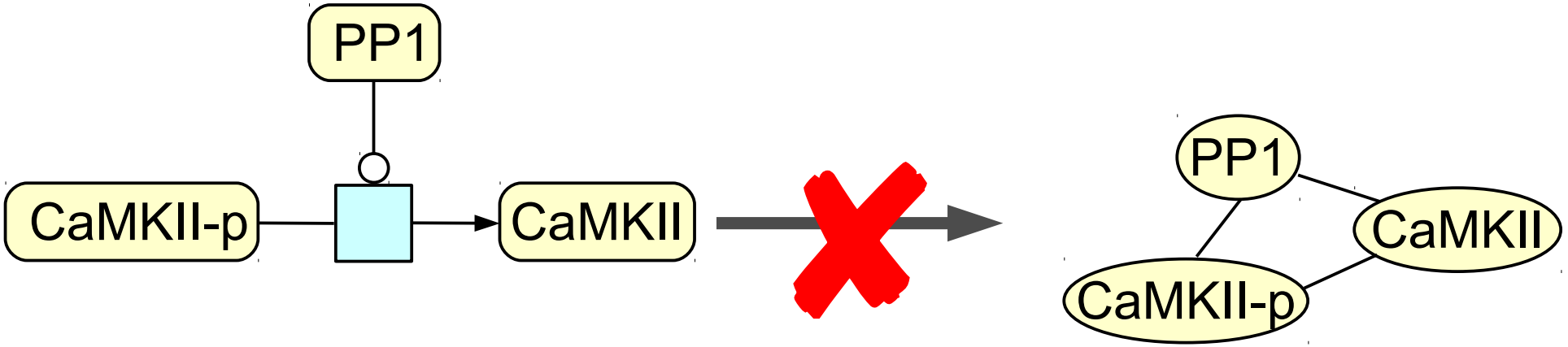
Useful insights at each level

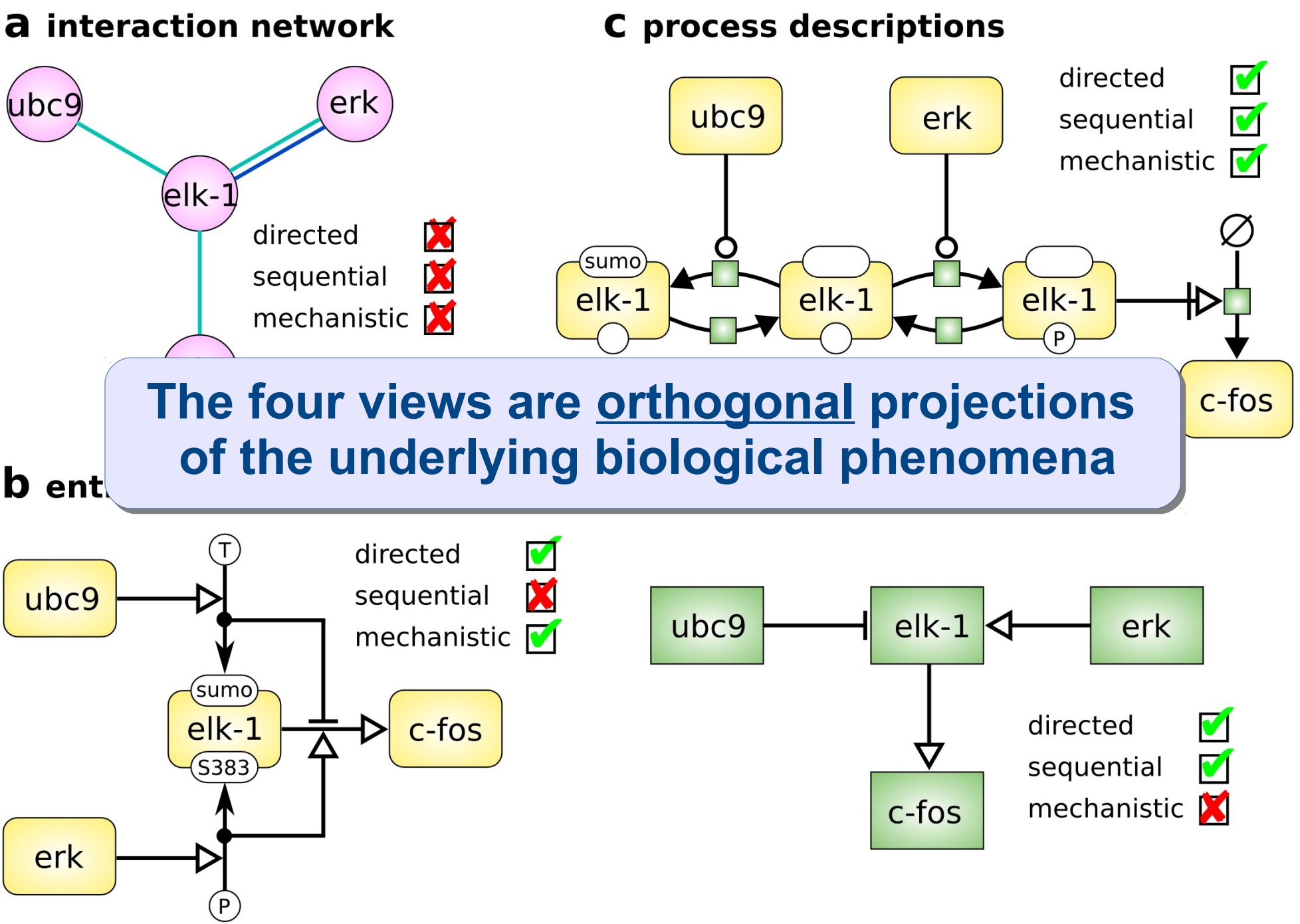
The four views of systems biology



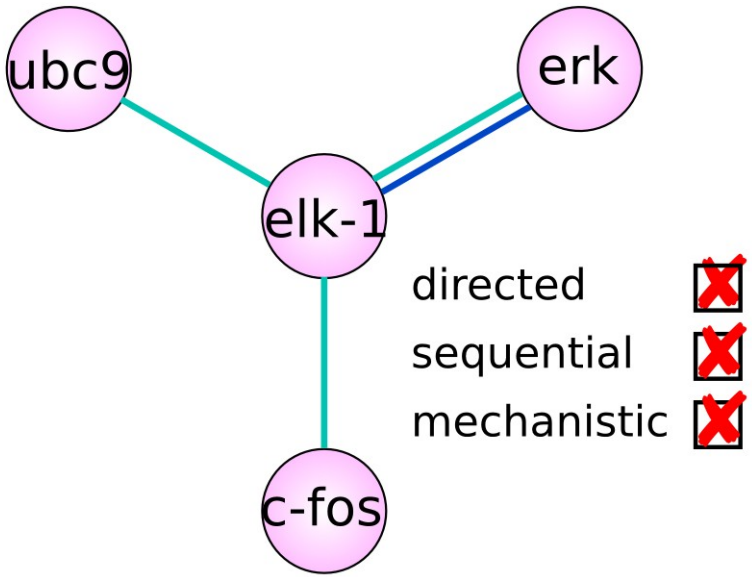

α-actinin
NMDAR
F-actin
PSD95
SPAR
PSD95
NMDAR


α-actinin
NMDAR
F-actin
PSD95
SPAR
PSD95
NMDAR


PP1
CaMKII-p
CaMKII
PP1
CaMKII
CaMKII-p


a interaction network
ubc9
erk
elk-1
directed
sequential
mechanistic
c process descriptions
ubc9
erk
directed
sequential
mechanistic
sumo
elk-1
elk-1
elk-1
P
c-fos
The four views are orthogonal projections of the underlying biological phenomena
b ent
T
ubc9
directed
sequential
mechanistic
sumo
elk-1
S383
c-fos
erk
P
ubc9
elk-1
erk
c-fos
directed
sequential
mechanistic

Interaction networks

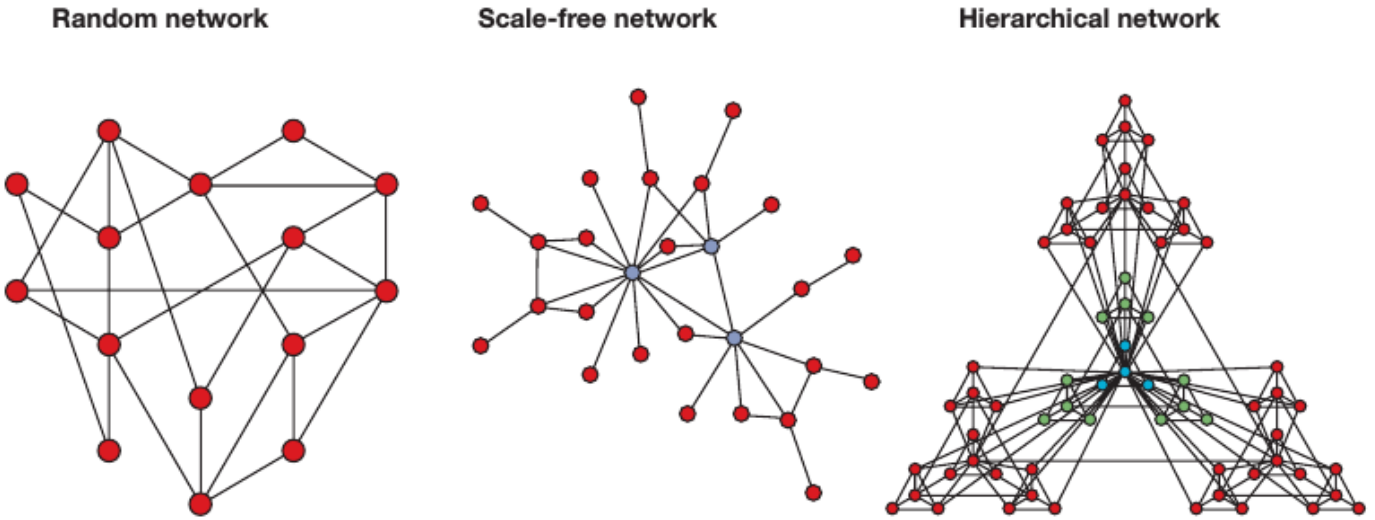


- Statistical modelling
- Functional genomics

What can we get from interaction networks

- Characteristics: degree of nodes (k) [k_{in}, k_{out} for directed networks], shortest and mean path lengths (ℓ), clustering coefficient (C)

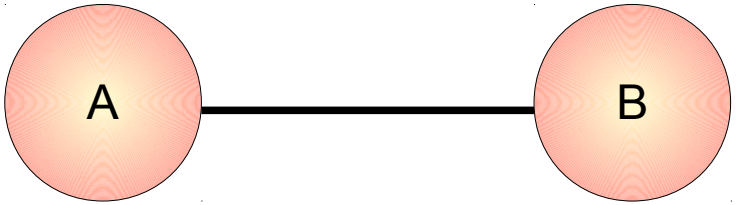
- Type of network



- Subgraphs, modules, motifs, motifs clusters …

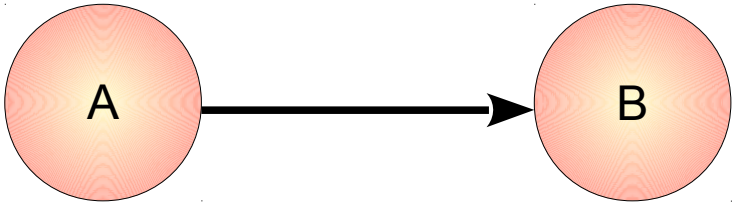
Barabasi *et al* (2004)

Undirected, directed, signed



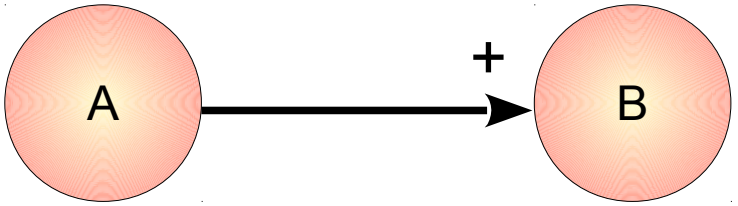
Undirected

"A interacts with B"



directed

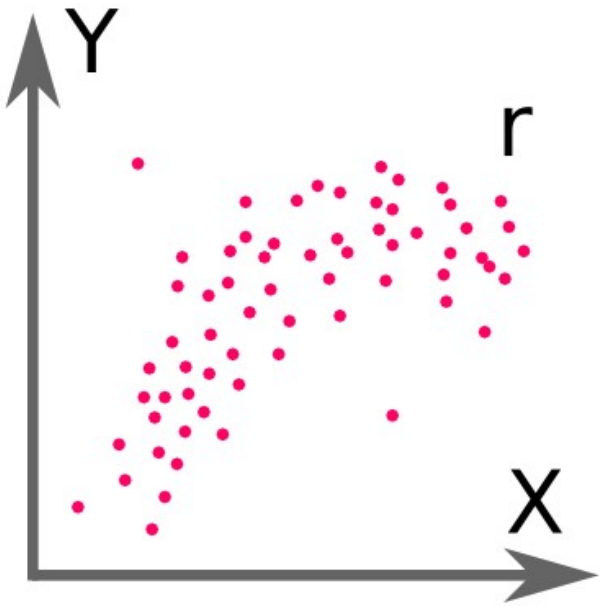
"A influences B"



Signed

A influences positively B

Network inference



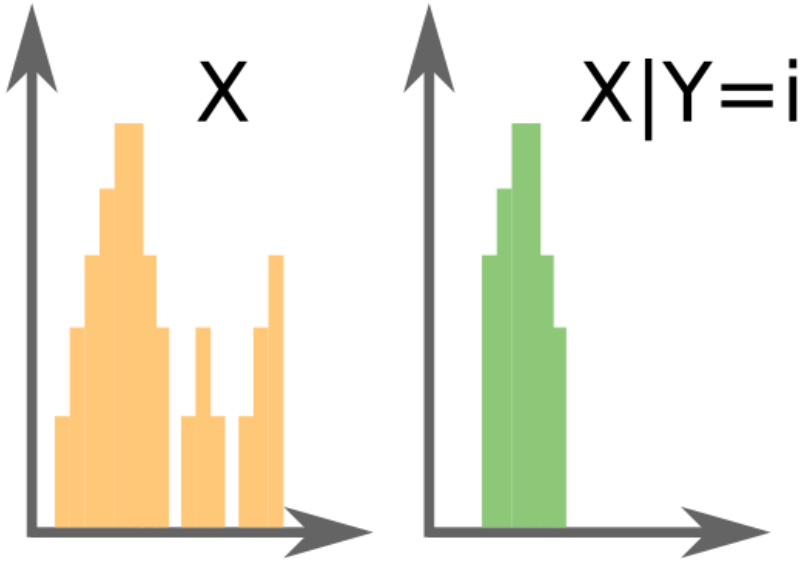
Correlation (e.g. regression)

Values of X are somewhat related to values of Y

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Villaverde *et al* (2014)

Network inference

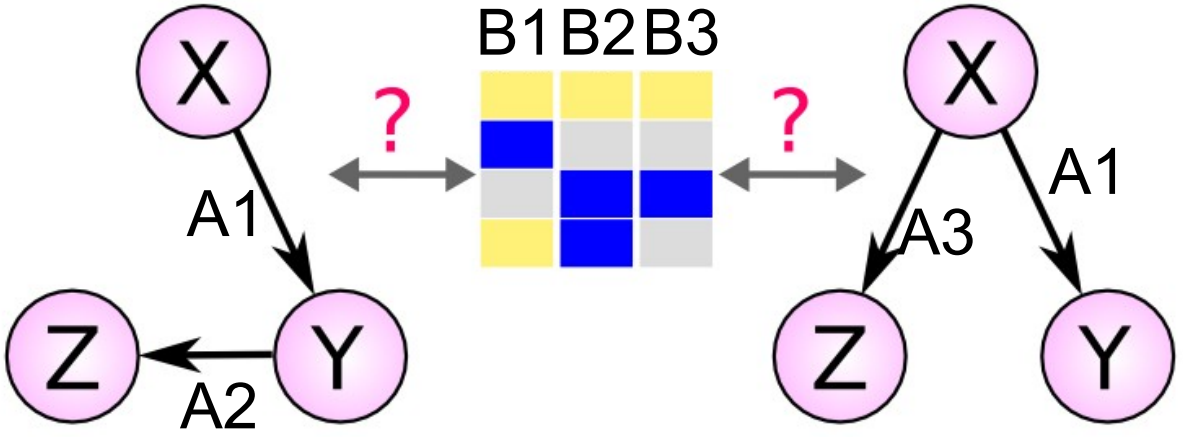


$$I(X,Y) = H(X) - H(X|Y)$$

Information theory

E.g. mutual information:
knowledge of the value of Y reduces the uncertainty on the values of X

Network inference

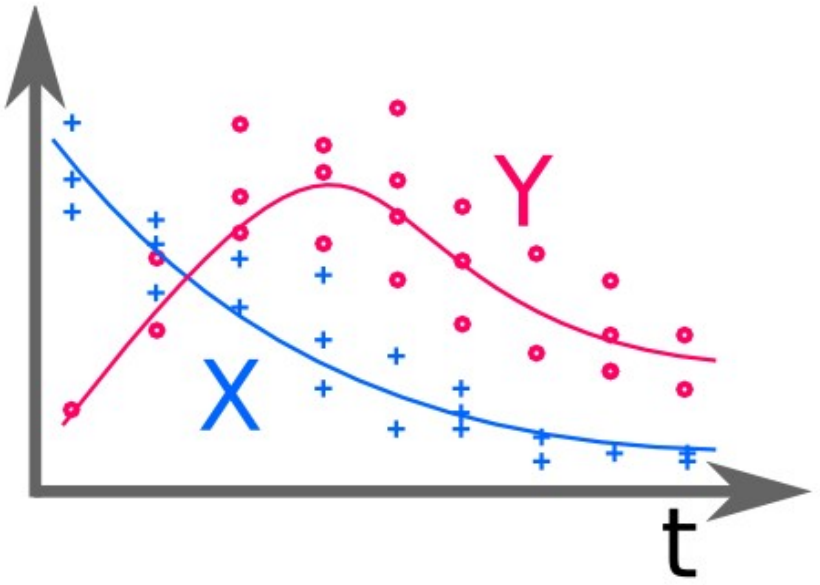


$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayesian inference

Which network most probably generates this dataset.

Network inference



ODEs

Equations describing time courses best

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{i,j} x_j$$

http://www.ebi.ac.uk/IntAct

Search: Q01097 Search Clear Show Advanced Fields » MIQL syntax reference

- Free text search will look by default for interactor identifier, species, interaction id, detection method, interaction type, publication identifier or author
- For a more specific search, use MIQL syntax or advanced search
- Search based on exact word matches eg. BRCA2 will not match BRCA2B
- Search for isoforms of 'P12345' by using 'P12345*'

- Home
- Advanced Search
- Tools
- Data Submission
- Downloads
- Documentation
- Acknowledgements
- Contact Us

FOLLOW US ON twitter

about 8 days ago IntAct release 158/159! 302465 binary interactions from 5787 manually curated publications ftp://ftp.ebi.ac.uk/pub/databases/intact/current

about 13 days ago IntAct is now available for download in PSI MITAB2.7 in addition to PSI-MI XML, MITAB2.5 and MITAB2.6.

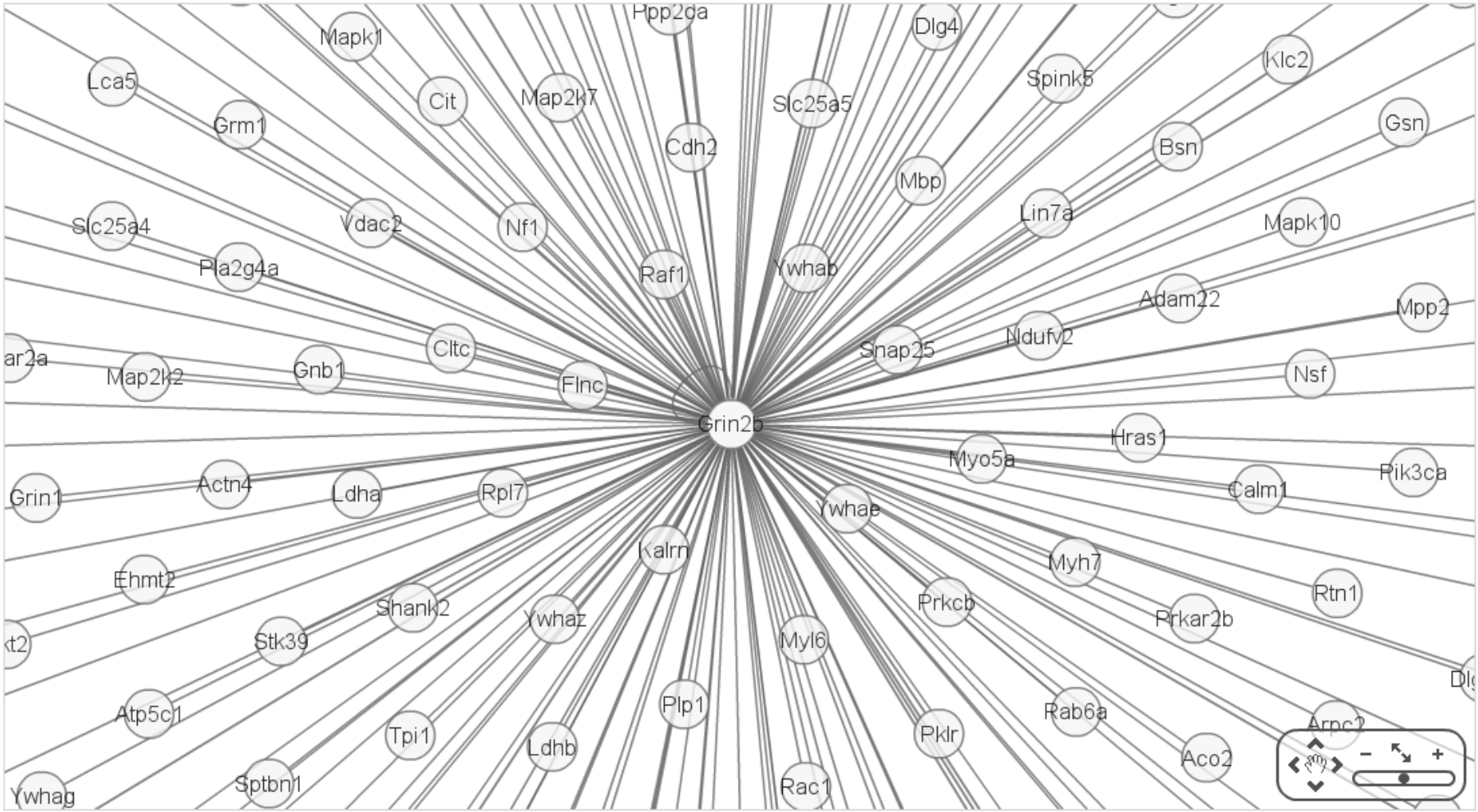
about 13 days ago Tarassov et al (PMID:18467557) low confidence interactors withdrawn in consultation with the authors

about 21 days ago Dataset of the month (November). Blandin et al. A Human Skeletal Muscle Interactome Centered on Proteins involved in Muscular Dystrophies.

about 34 days ago Now IntAct can

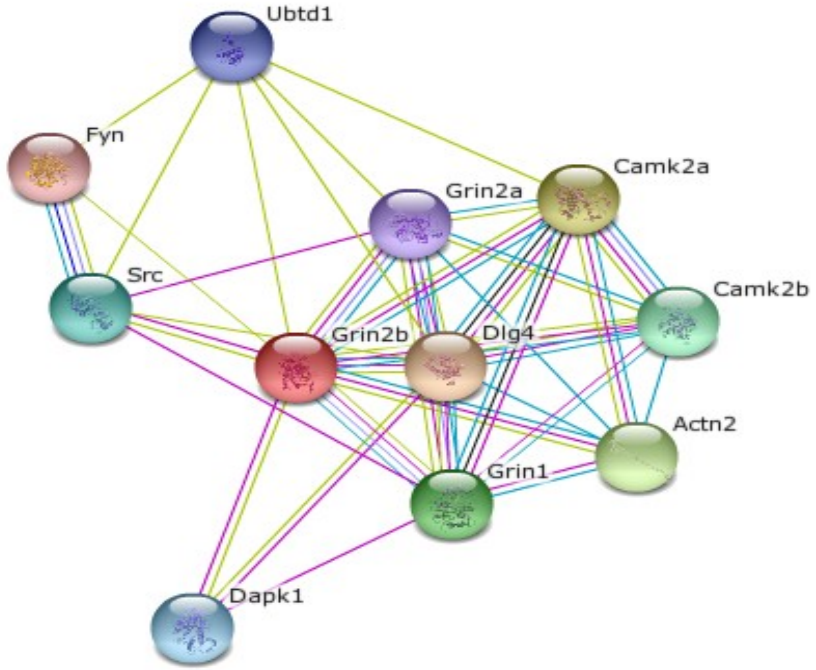
Home | Search | Interactions (254) | Browse | Lists | Interaction Details | Molecule View | Graph

Network visualisation



http://string-db.org/

Home · Download · Help/Info STRING 9.0



This is the **evidence view**. Different line colors represent the types of evidence for the association.

confidence evidence actions interactive advanced more less save

(requires Flash player 10 or better)

Your Input:

Grin2b glutamate receptor, ionotropic, NMDA2B (epsilon 2) Gene; NMDA receptor subtype of glutamate-gated ion channels with high calcium permeability and voltage-dependent sensitivity to magnesium. Mediated by glycine (1482 aa)
(Mus musculus)

Predicted Functional Partners:

Neighborhood Gene Fusion Cooccurrence Coexpression Experiments Databases Textmining [Homology] Score

Babraham Institute

Practical, tomorrow

Lu Li

Bhupinder Virk

The Systems Biology Graphical Notation

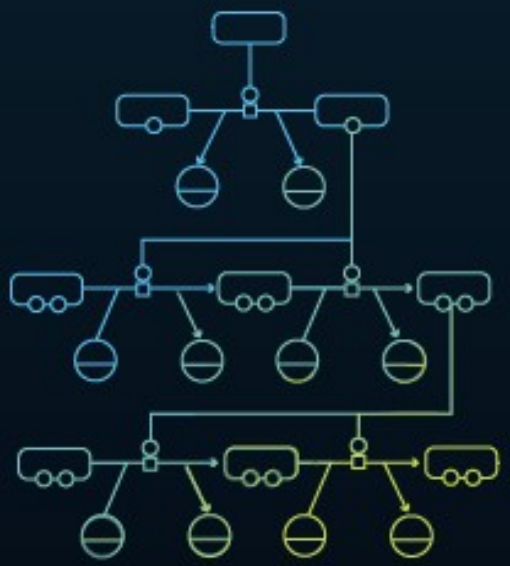
http://sbgn.org/

Le Novère *et al* (2009)

Unambiguous consensual visual notation

- An unambiguous way of graphically describing and interpreting biochemical and cellular events
- Limited amount of symbols
  Re-use existing symbols

  ☞ Smooth learning curve
- Can represent logical or mechanistic models, biochemical pathways, at different levels of granularity
- Detailed technical specification, precise data-models, standard API and growing software support
- Developed over ten years by a diverse community, including biologists, modellers, computer scientists etc.



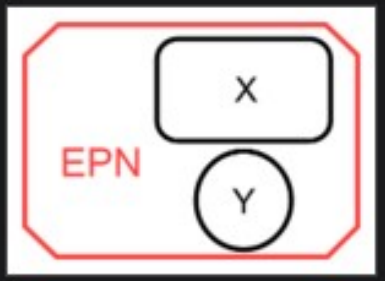
A Visual Notation for Network Diagrams in Biology

Welcome to the global portal for documentation, news, and other information about the **Systems Biology Graphical Notation** (SBGN) project, an effort to standardize the graphical notation used in maps of biological processes.

Quick start

Learn how to use SBGN	Get involved

Symbol of the month

Complex (PD)	

SBGN News

(19 Aug. '14) Draft versions of the latest SBGN specs are now available for review. Draft PD Level 1 Version 2.0 and draft ER Level 1 Version 2.0

Pathway of the month

This month's pathway is an SBGN-PD diagram based on the following paper published in the October 2014 issue of the Cell magazine (PubMed ID: 25303525).

Ma H, Groth RD, Cohen SM, Emery JF, Li B, Hoedt E, Zhang G, Neubert TA, Tsien RW. (2014) γCaMKII shuttles Ca^{2+}/CaM to the nucleus to trigger CREB phosphorylation and

The Systems Biology Markup Language

http://sbml.org/

Hucka *et al* (2003)

SBML supporting tools

- LibSBML: free, open-source programming library to help you read, write, manipulate, translate, and validate SBML files. Written in C++, with bindings for: c#, Python, Java, Perl, Ruby, MatLab, Octave
- JSBML: free, open-source, pure Java library for reading, writing, and manipulating SBML files (validation done via libSBML)
- SBML converters: Converter from and to SBML, including Octave, XPP, BioPAX, dot, SVG, MDL
- Software guide: >250 software, including modelling and simulation environment, databases, model processing

- Core package – public specification → Process Descriptions
- Flux balance constraint – public specification
- Qualitative models – public specification → Activity Flows
- Model composition – public specification
- Graph Layout – public specification
- Graph rendering – specification finalised
- Complex species – specification finalised → Entity Relationships
- Groups - specification finalised
- Spatial diffusion – specification under discussion
- Distributions and ranges - specification under discussion
- Enhanced metadata – specification proposed
- Arrays and sets – specification proposed
- Dynamic structures - discussed

SBML Level 3 is modular

SBML.org — The Systems Biology Markup Language

Welcome to the portal for the **Systems Biology Markup Language** (**SBML**), a free and open interchange format for computer models of biological processes. SBML is useful for models of metabolism, cell signaling, and more. It **continues to be evolved and expanded** by an international community.

For the curious

What *is* SBML? Read our **introduction**, then perhaps browse the **mailing lists**, the **FAQ**, and look at the **SBML Level 3 package activities** to glimpse what's happening with SBML today.

For modelers

Looking for software that supports SBML? Our **software guide** lists over **250** systems. Are you instead looking for models? Visit **BioModels Database**, where you can find hundreds!

For software developers

Want to support SBML in your software? Read our **intro** and then the **specifications** to understand SBML in depth, then check our **libraries**, **test resources**, and also **3rd-party software**.

No matter how you use SBML, we invite you to sign up for news updates either through our **RSS feed**, our **Twitter feed**, or one of the **mailing lists**, and get involved with **community efforts** to help keep improving SBML. You can also call attention to your project's support of SBML by displaying the **SBML logo**.

SBML would not have been possible without support from **multiple agencies and organizations**, as well as intellectual contributions from many motivated individuals, including the **major contributors** who are shaping SBML Level 3.

SBML News

SBML Layout spec. released

(14 Aug.'13) The SBML Level 3 Layout V.1 specification has been finalized and released.

COMBINE 2013

(28 Jul.'13) Registration and a preliminary agenda for the September event are now available.

SBML Test Runner 3.0

(6 Jun.'13) A new version of the standalone Test Runner for the SBML Test Suite is now available.

Older news ...

Community News

SBML2GPML

(12 Sep.'13) This new PathVisio plugin allows SBML models to be imported.

CellNOpt 1.7.7

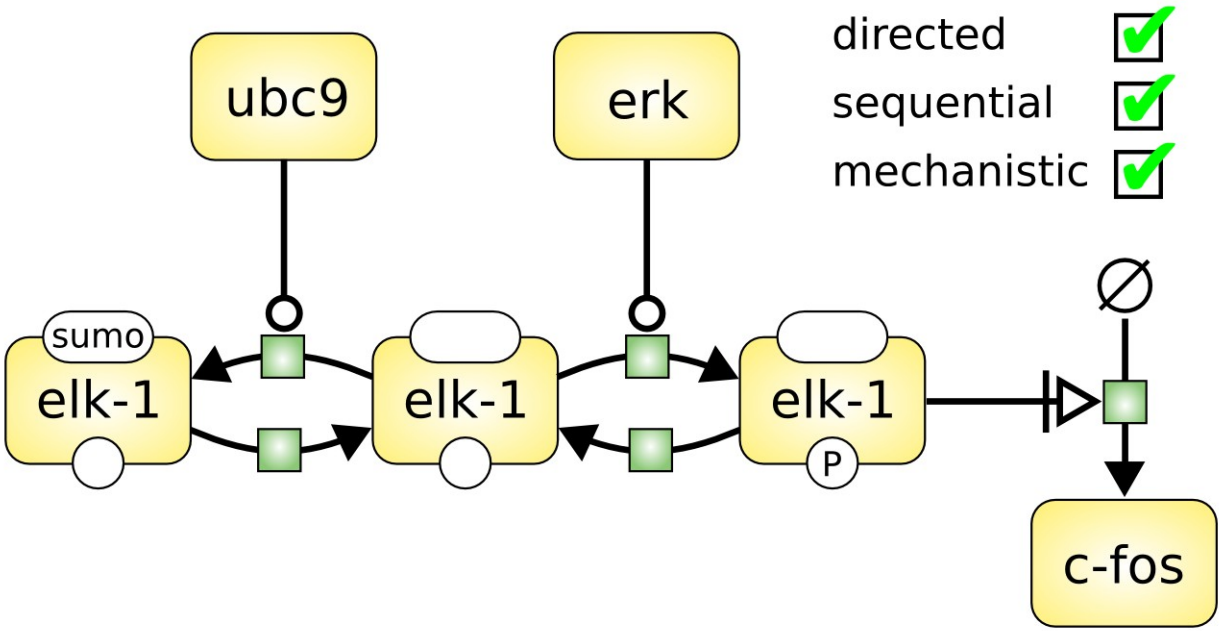
(12 Aug.'13) This logic-based modeling software system now supports SBML Level 3 Qualitative Models.

iNA 0.4.2 released

(11 Jun.'13) The new release offers a new plot editor and supports multi-core computations on MacOS.

Older news ...

Process Descriptions



- Process modelling
- Biochemistry, Metabolic networks
- Generally within “closed world”
- Subjected to combinatorial explosion

Open world

Anything not explicitly stated is unknown

Failure to observe does not imply non-existence

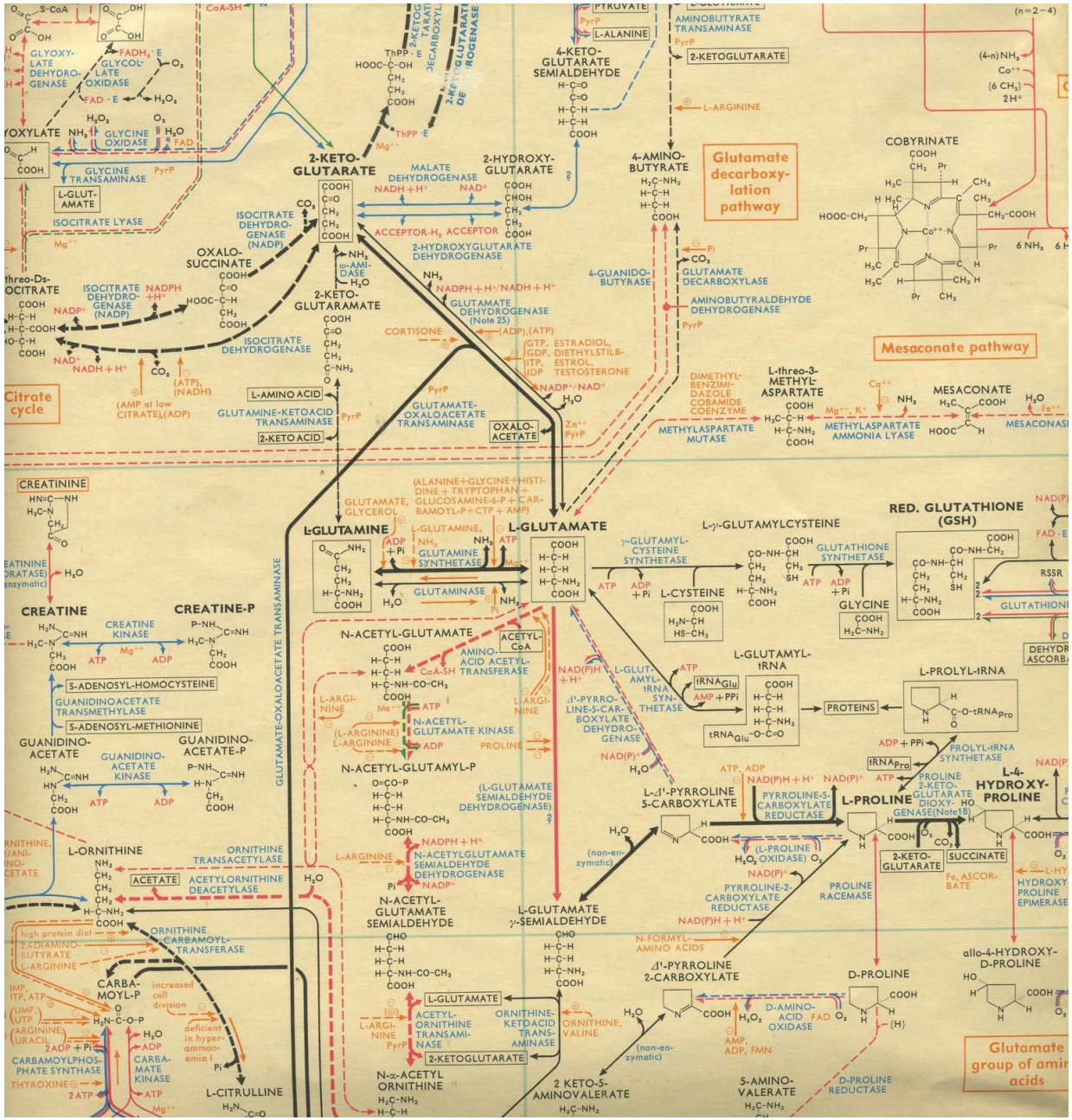
New pieces of knowledge do not affect prior pieces

Closed world

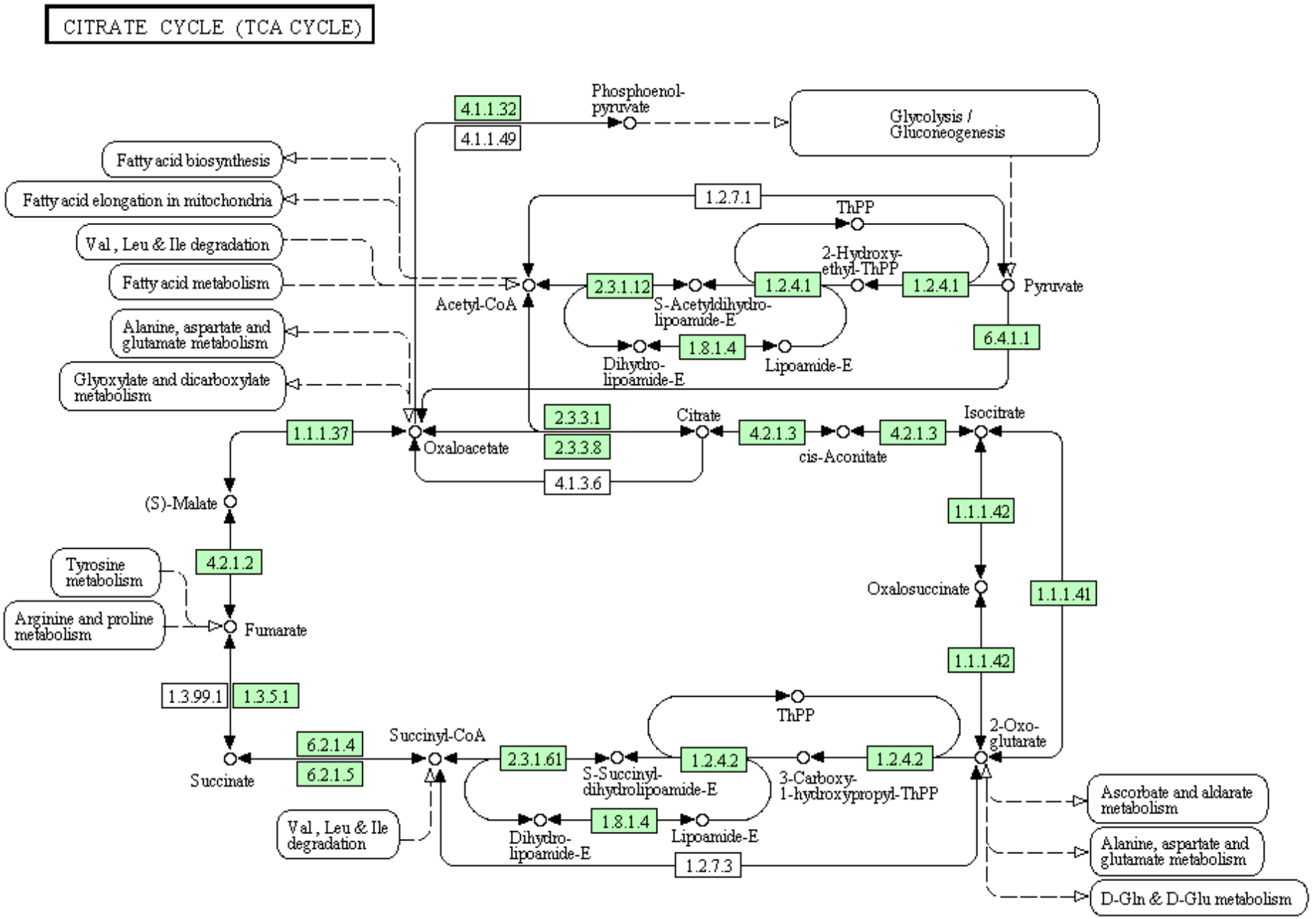
Anything not explicitly stated does not exist

Failure to observe implies non-existence

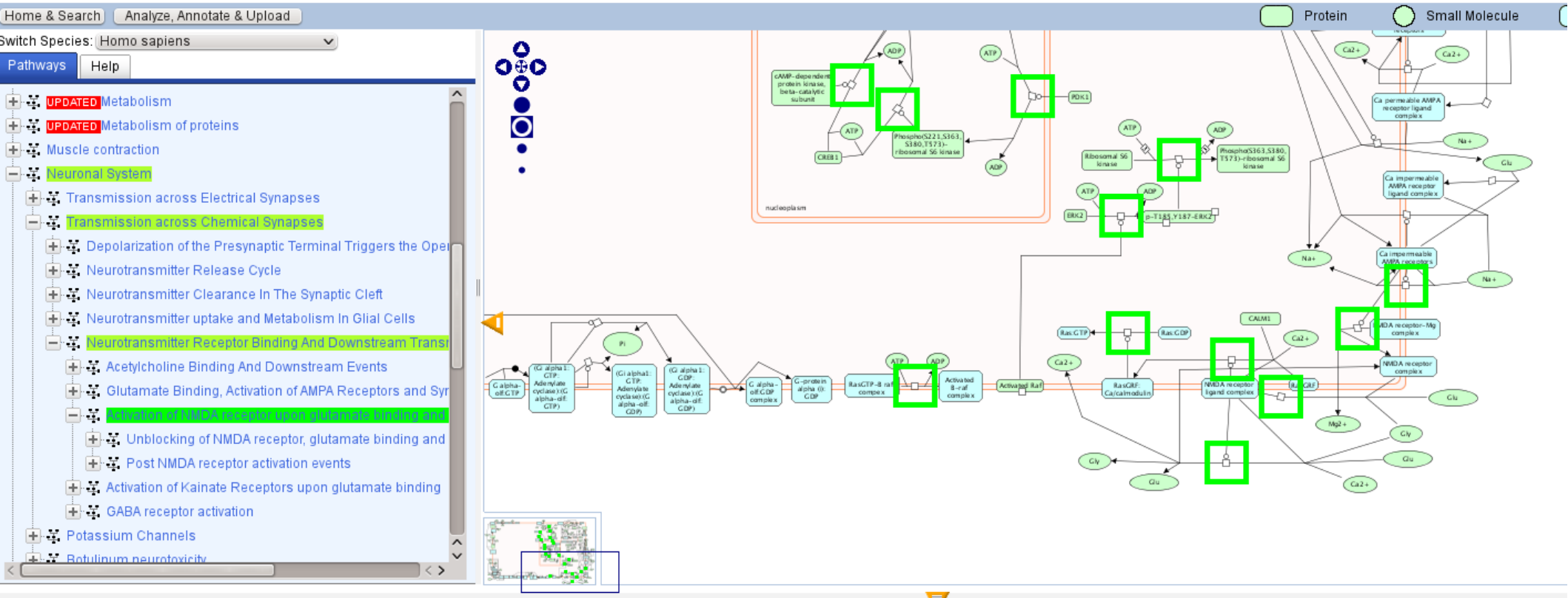
New pieces of knowledge might change the meaning of prior pieces


2-KETO-GLUTARATE
2-HYDROXY-GLUTARATE
4-AMINO-BUTYRATE
Glutamate decarboxy-lation pathway
COBYRINATE
Citrate cycle
OXALO-SUCCINATE
2-KETO-GLUTARAMATE
Mesaconate pathway
L-GLUTAMINE
L-GLUTAMATE
L-γ-GLUTAMYLCYSTEINE
RED. GLUTATHIONE (GSH)
CREATININE
CREATINE
CREATINE-P
N-ACETYL-GLUTAMATE
L-GLUTAMYL-tRNA
L-PROLYL-tRNA
GUANIDINO-ACETATE
GUANIDINO-ACETATE-P
L-ORNITHINE
N-ACETYL-GLUTAMYL-P
L-Δ¹-PYRROLINE 5-CARBOXYLATE
L-PROLINE
L-4-HYDROXY-PROLINE
N-ACETYL-GLUTAMATE SEMIALDEHYDE
L-GLUTAMATE γ-SEMIALDEHYDE
Δ¹-PYRROLINE 2-CARBOXYLATE
D-PROLINE
allo-4-HYDROXY-D-PROLINE
L-CITRULLINE
N-α-ACETYL ORNITHINE
2 KETO-5-AMINOVALERATE
5-AMINO-VALERATE
Glutamate group of amino acids

http://www.genome.jp/kegg/pathway.html



http://www.reactome.org



Activation of NMDA receptor upon glutamate binding and postsynaptic events

Stable identifier	REACT_20563.2
Authored	Mahajan, SS, 2009-10-29
Reviewed	Tukey, D, 2009-11-17

NMDA receptors are a subtype of ionotropic glutamate receptors that are specifically activated by a glutamate agonist N-methyl-D-aspartate (NMDA). Activation of NMDA receptor involves opening of the ion channel that allows the influx of Ca2+. NMDA receptors synaptic strength and are predominantly involved in the synaptic plasticity that pertain to learning and memory. A unique feature of NMDA receptor unlike other glutamate receptors is the requirement of dual activation of the NMDA receptor, which require bo activation. At resting membrane potential the NMDA receptors are blocked by Mg2+. The voltage dependent Mg2+ block is relieved upon depolarization of the post-synpatic membrane . The ligand dependent activation of the NMDA receptor requires co-activ glycine. NMDA receptors are coincidence detector, and are activated only if there is simultaneous activation of both pre and post-synaptic cell. Upon activation NMDA receptors allow the influx of Ca2+ that initiates various molecular signaling cascades that memory.

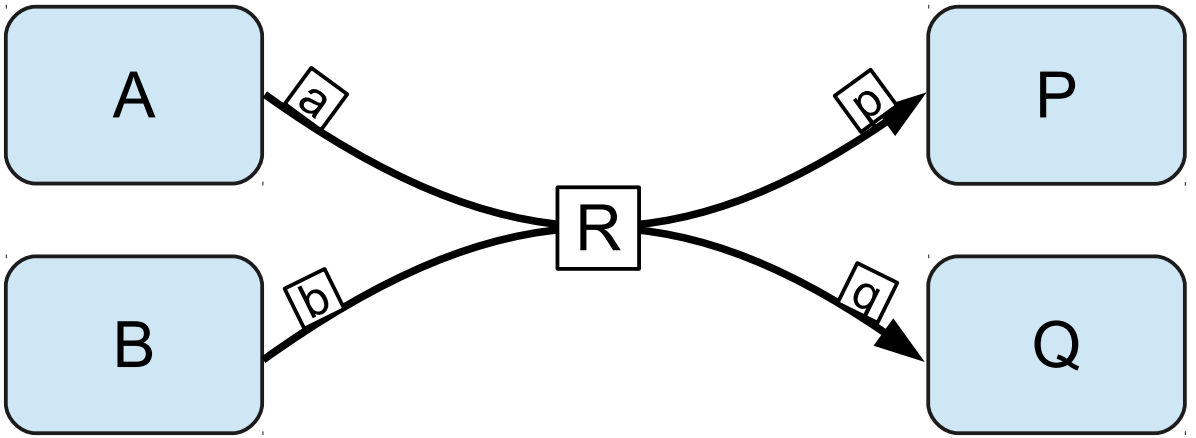
Organism	Homo sapiens
Cellular compartment	extracellular region GO plasma membrane GO cytoplasm GO nucleoplasm GO

References

Cohen, S, Greenberg, ME *Communication between the synapse and the nucleus in neuronal development, plasticity, and disease* 2008 Annu Rev Cell Dev Biol PubMed

	Activation of NMDA receptor upon glutamate binding and postsynaptic events [Dictyostelium discoideum] Activation of NMDA receptor upon glutamate binding and postsynaptic events [Schizosaccharomyces pombe] Activation of NMDA receptor upon glutamate binding and postsynaptic events [Saccharomyces cerevisiae] Activation of NMDA receptor upon glutamate binding and postsynaptic events [Caenorhabditis elegans]

A biochemical reaction is a process

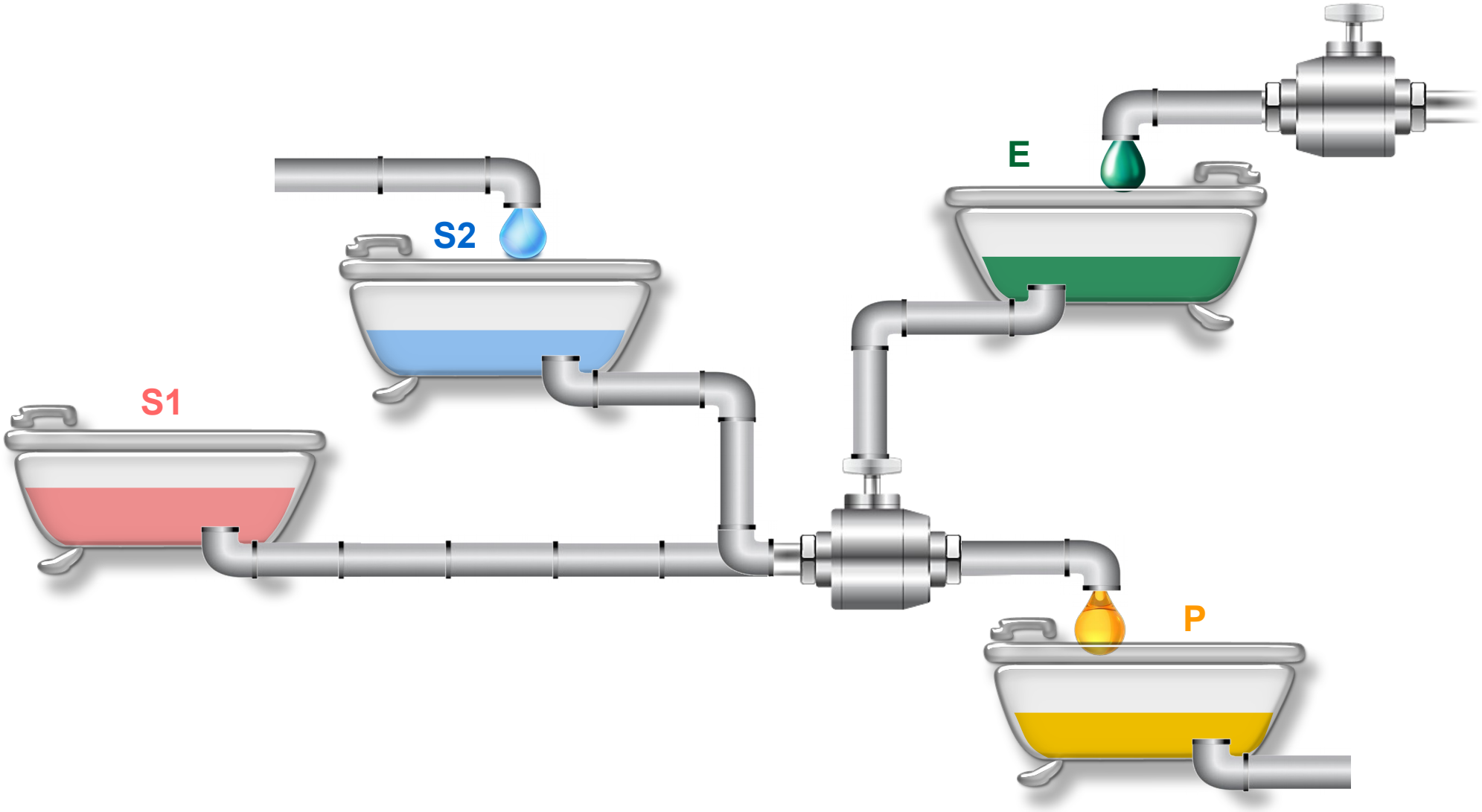


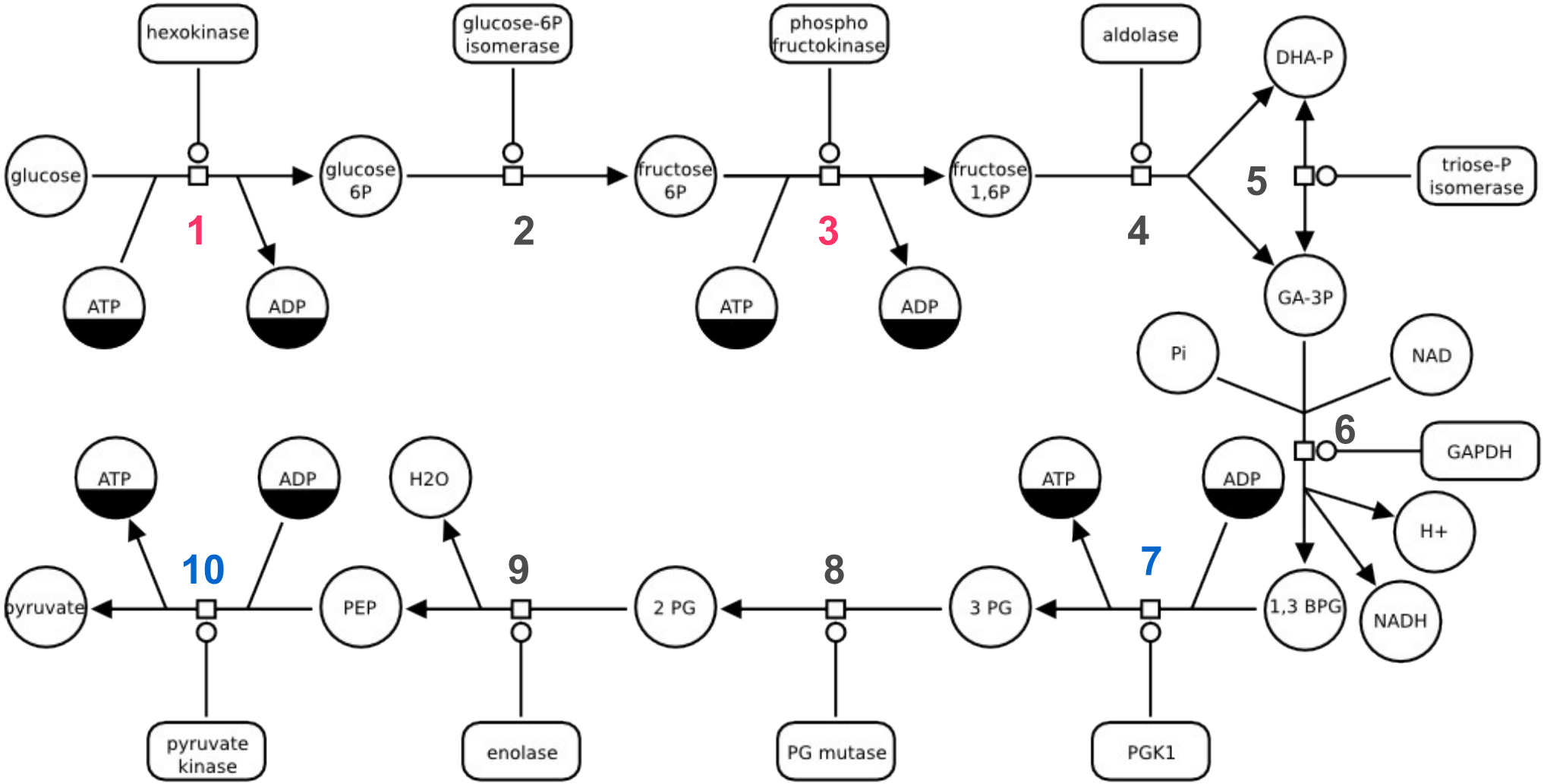
substances A and B are consumed by reaction R that produces substances P and Q

→ Reconstruction of state variable evolution from process descriptions:

- Processes can be combined in ODEs (for deterministic simulations); transformed in propensities (for stochastic simulations)
- Systems can be reconfigured quickly by adding or removing a process

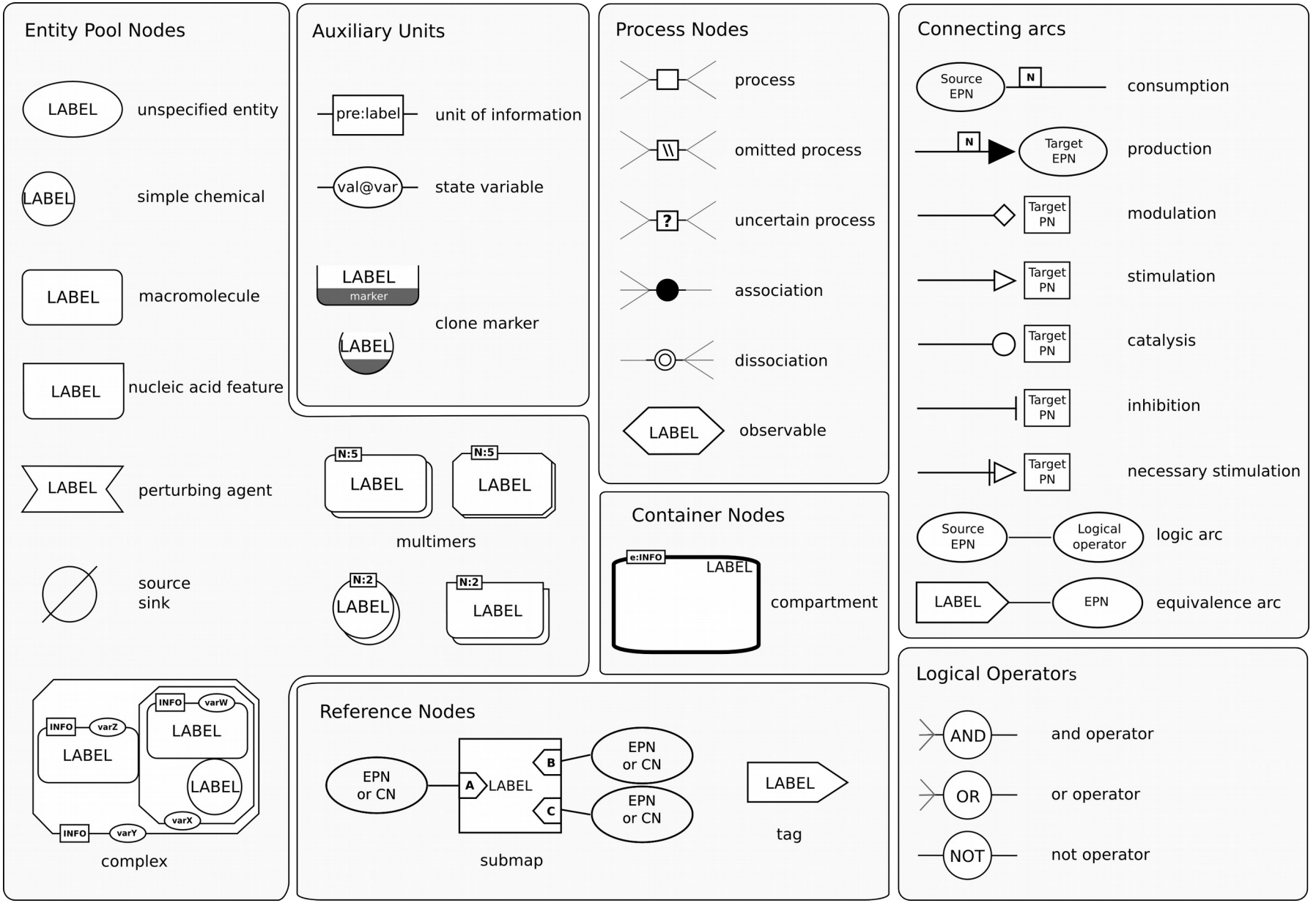
Process Descriptions can be viewed as pipelines



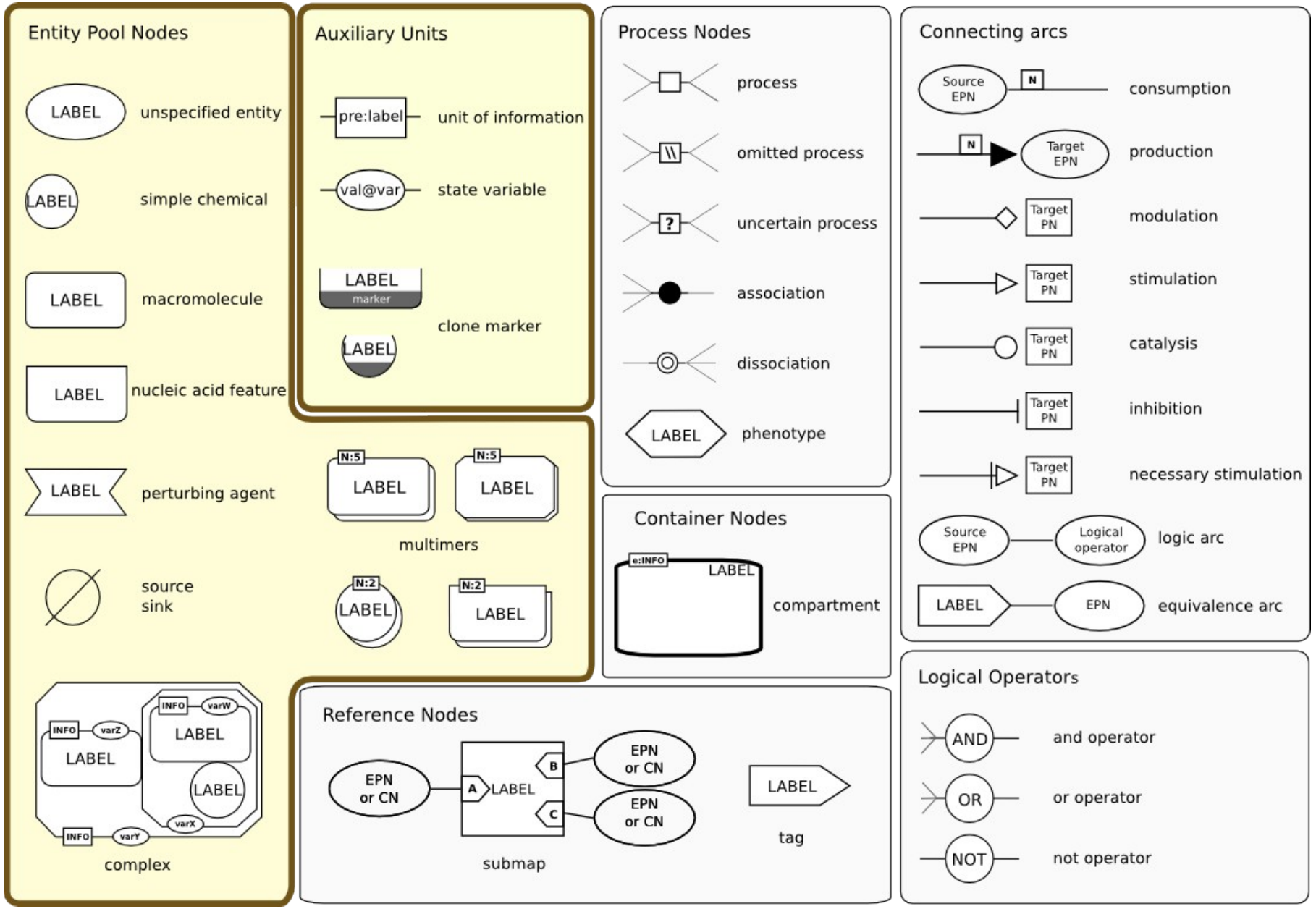


ATP is consumed by processes 1 and 3, and produced by processes 7 and 10

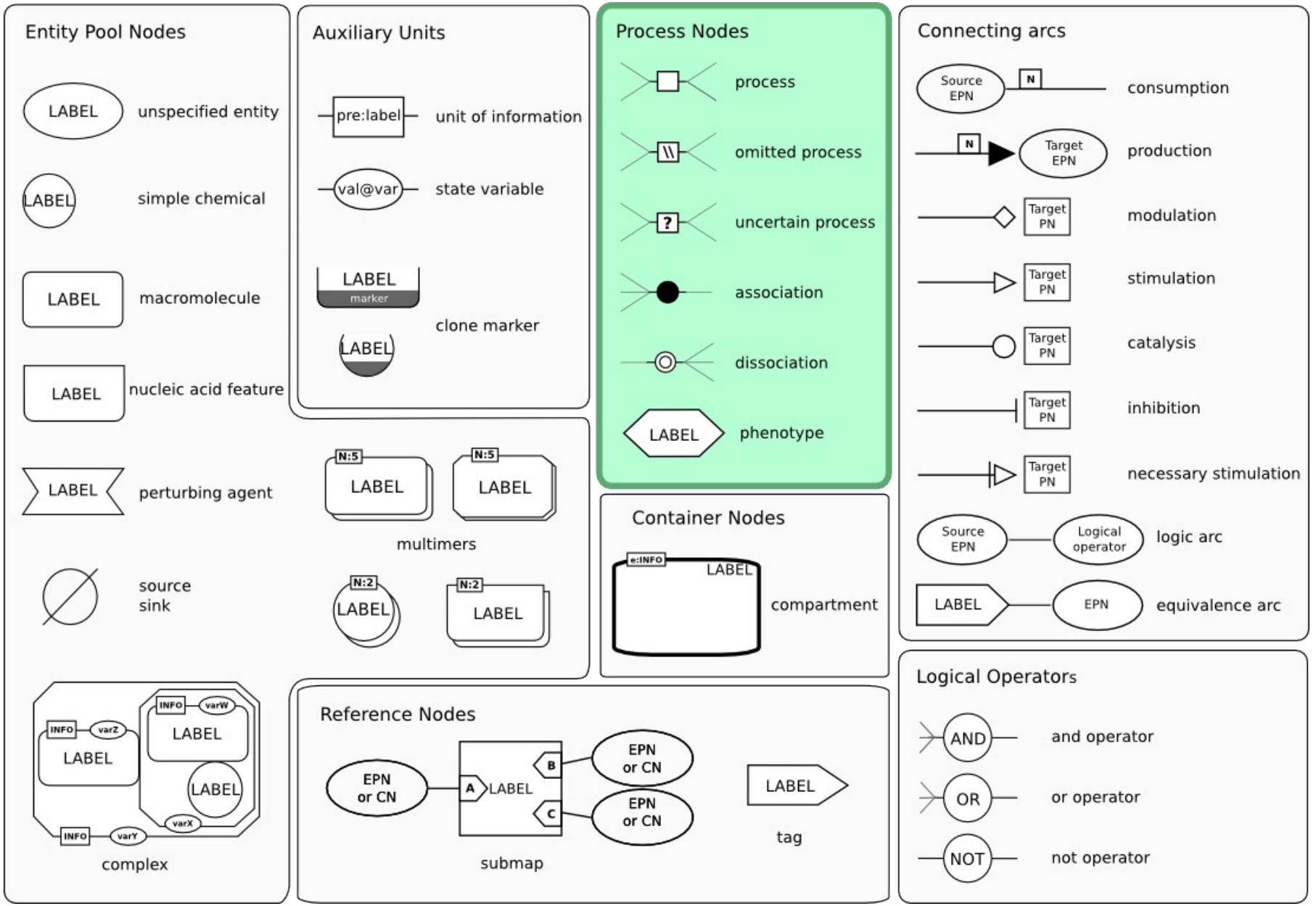
SBGN Process Diagram L1 reference card



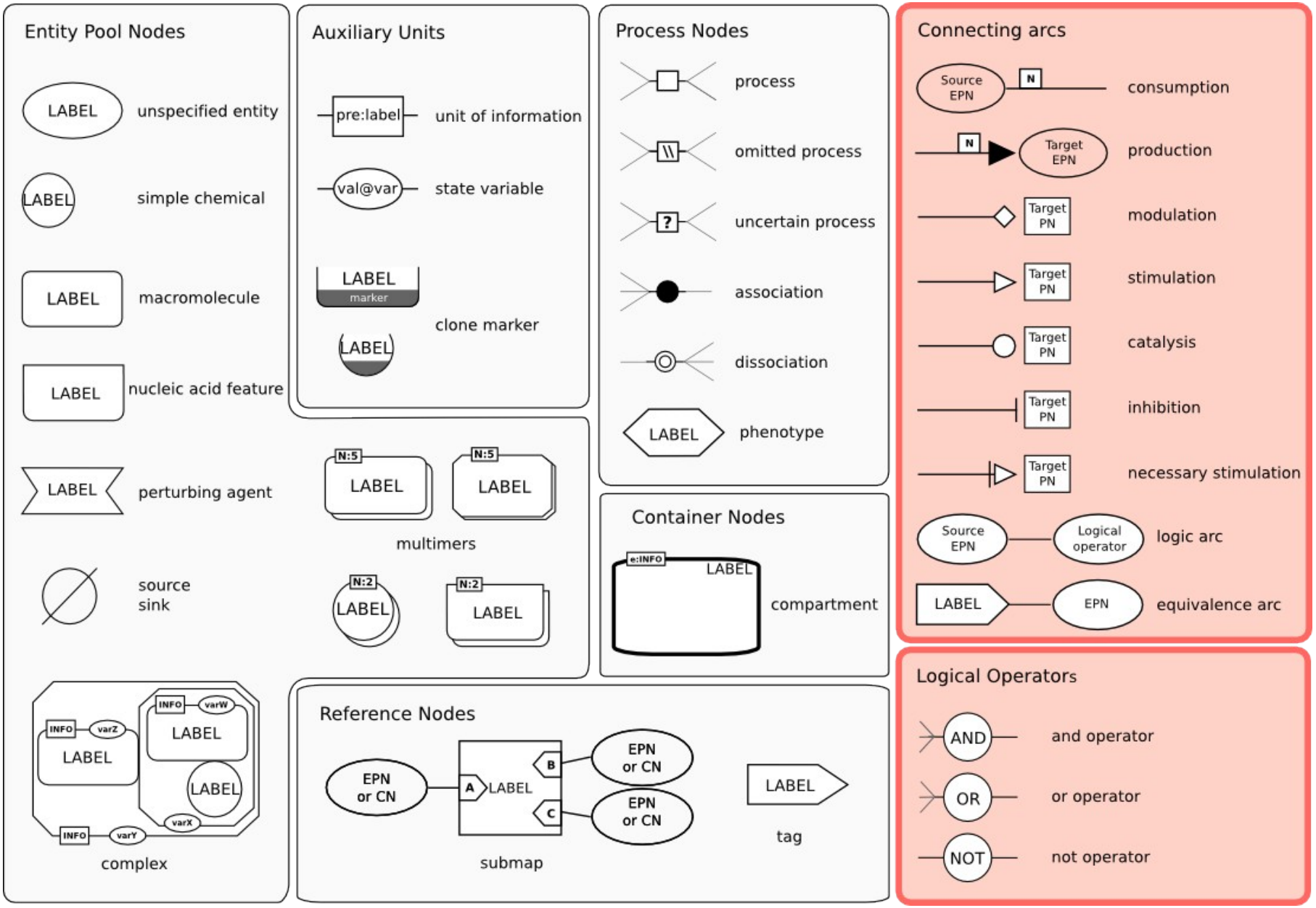
Entity Pool Nodes



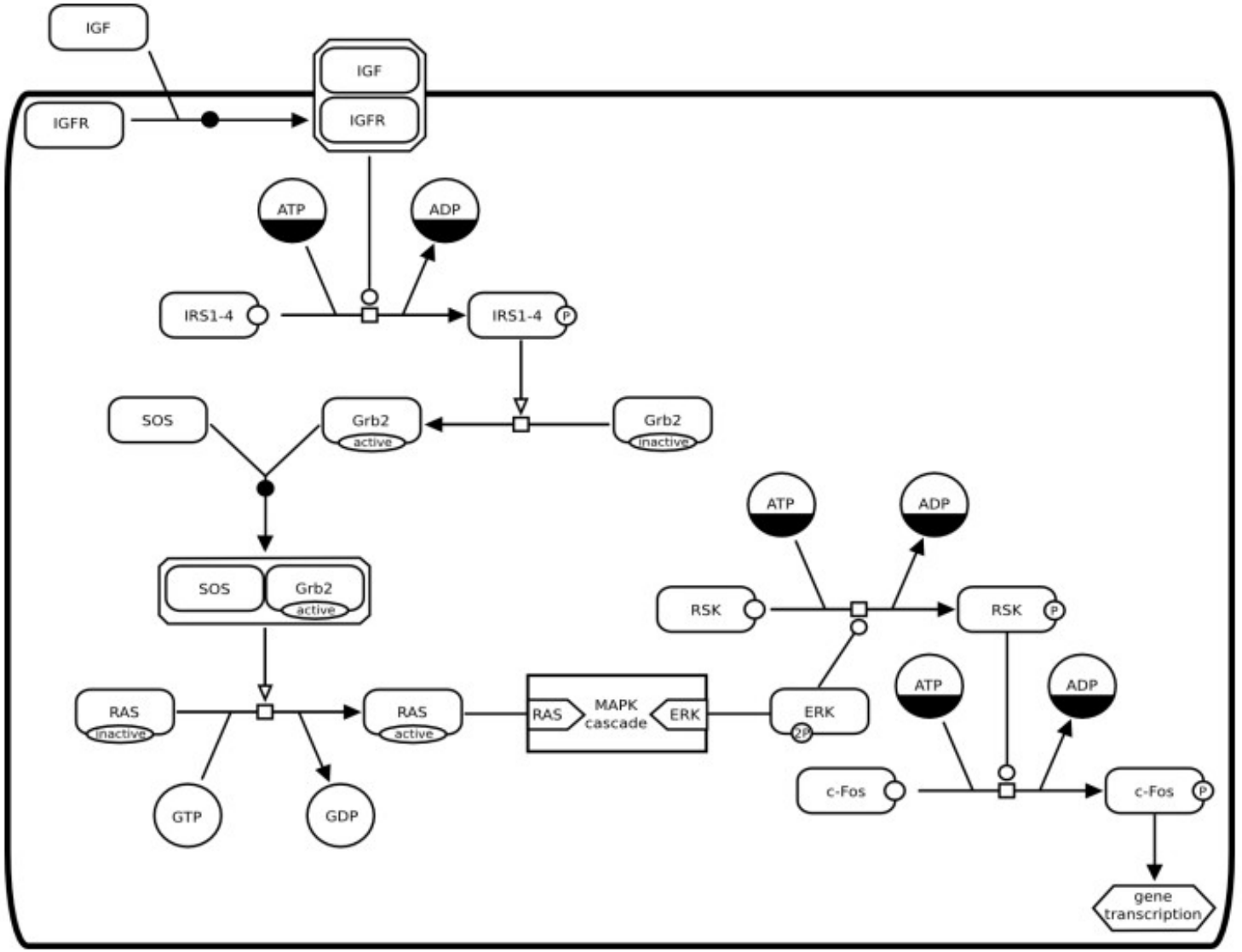
Process Nodes



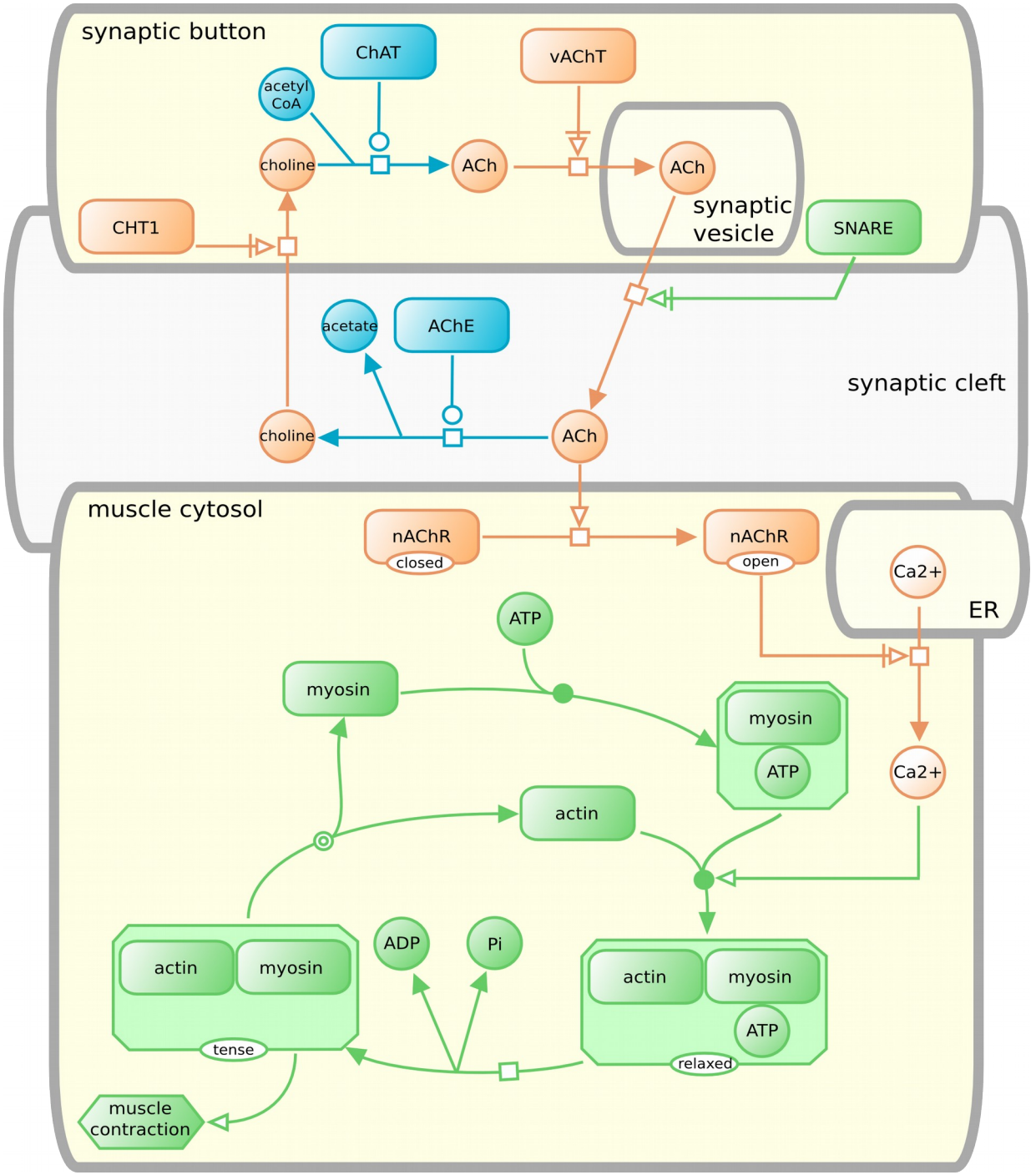
Connecting arcs

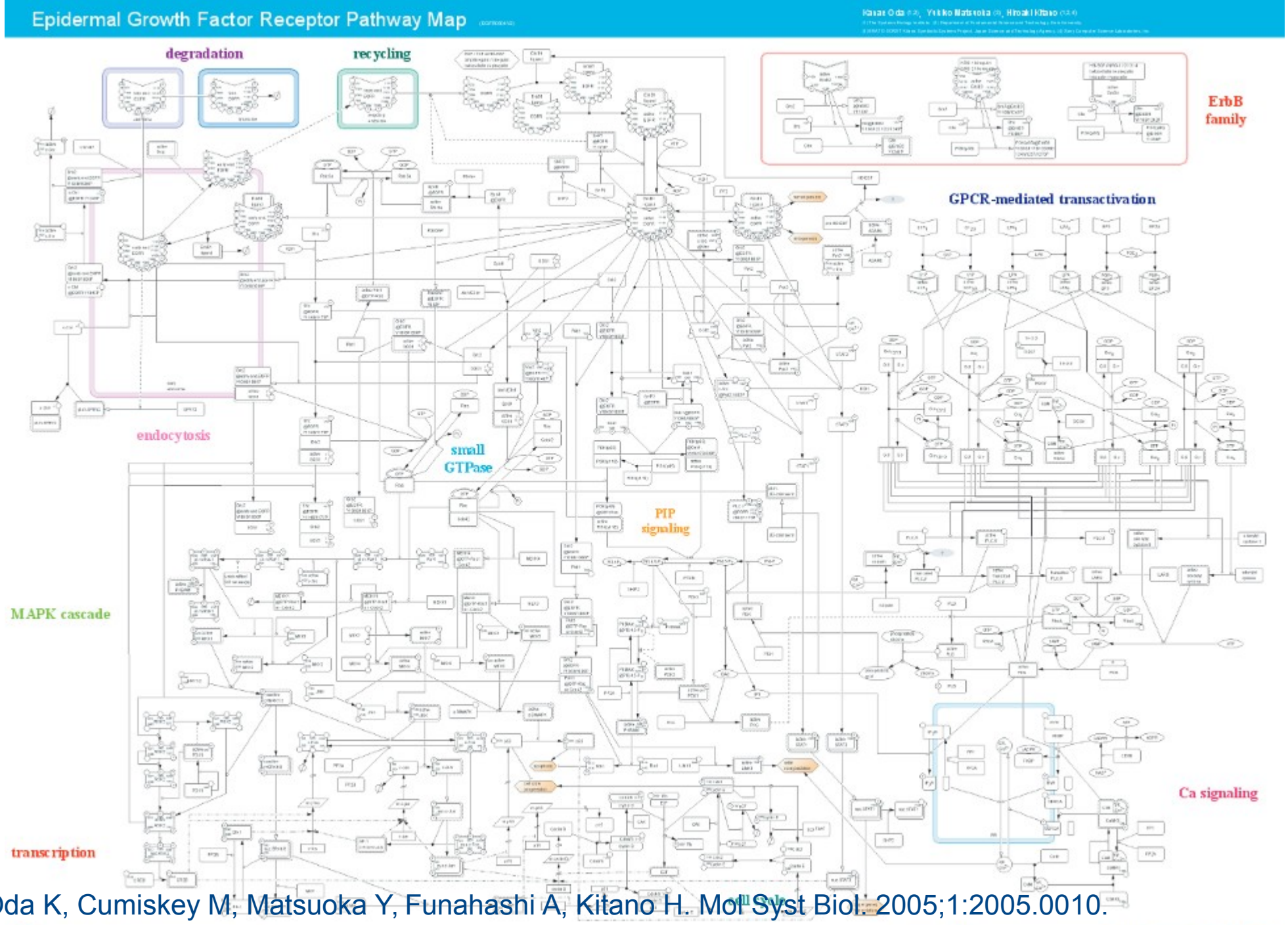


Signalling

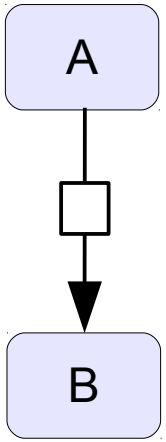

IGF
IGFR
IGF
IGFR
ATP
ADP
IRS1-4
IRS1-4
P
SOS
Grb2
active
Grb2
inactive
SOS
Grb2
active
ATP
ADP
RSK
RSK
P
RAS
inactive
RAS
active
RAS
MAPK
cascade
ERK
ERK
2P
ATP
ADP
GTP
GDP
c-Fos
c-Fos
P
gene
transcription

Multicellular





Oda K, Cumiskey M, Matsuoka Y, Funahashi A, Kitano H. Mol Syst Biol. 2005;1:2005.0010.

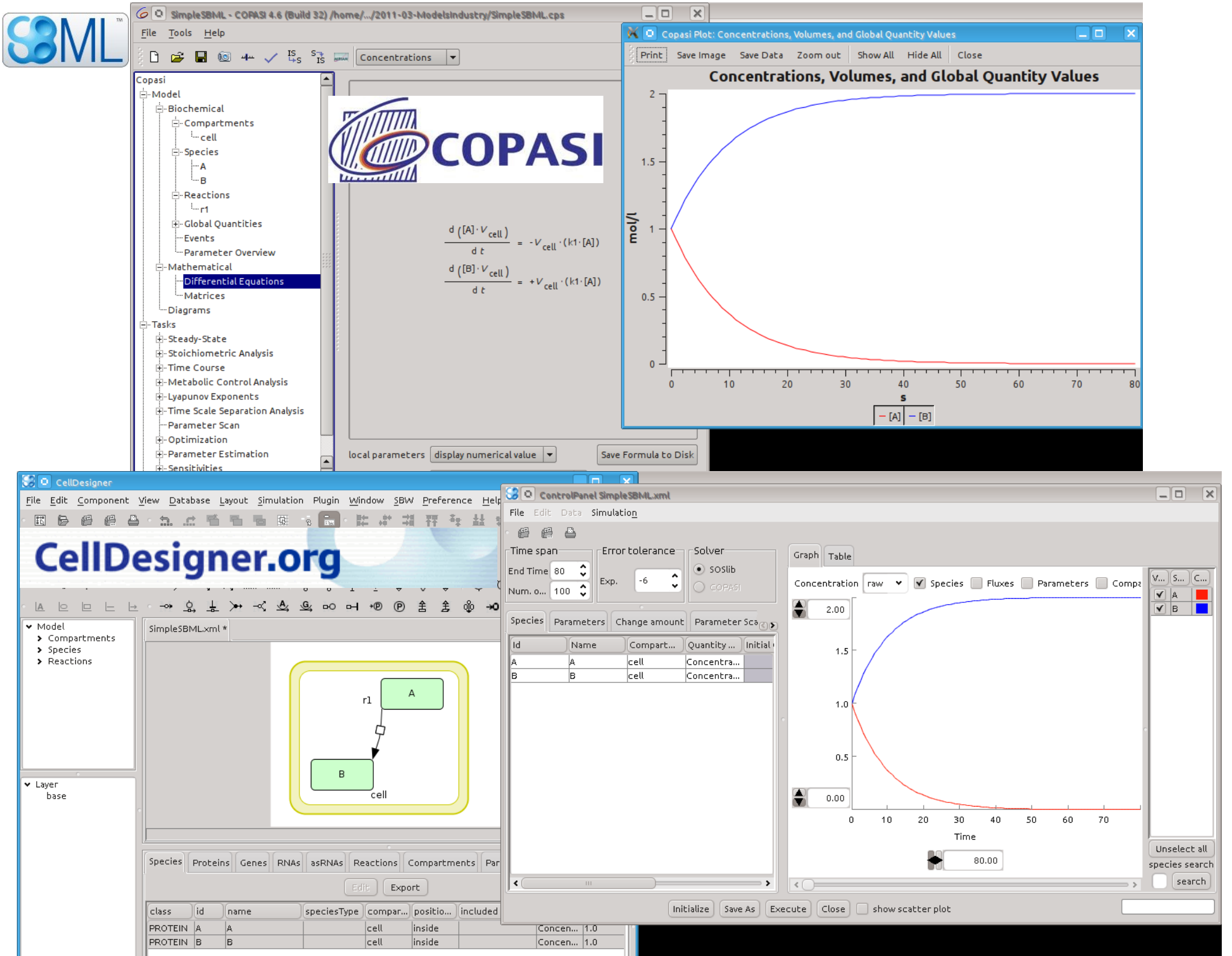


$$\frac{d[B]}{dt} = k \times [A]$$

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2/version4" level="2" version="4">
  <model name="Simple Model">
    <listOfCompartments>
      <compartment id="cell" size="1" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="1"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="k1" value="0.1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction id="r1" reversible="false">
      <listOfReactants>
          <speciesReference species="A"/>
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B"/>
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times/>
              <ci> cell </ci>
              <ci> k1 </ci>
              <ci> A </ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

A very simple SBML file

http://sbml.org


SBML
SimpleSBML - COPASI 4.6 (Build 32) /home/.../2011-03-ModelsIndustry/SimpleSBML.cps
File Tools Help
Concentrations
Copasi
Model
Biochemical
Compartments
cell
Species
A
B
Reactions
r1
Global Quantities
Events
Parameter Overview
Mathematical
Differential Equations
Matrices
Diagrams
Tasks
Steady-State
Stoichiometric Analysis
Time Course
Metabolic Control Analysis
Lyapunov Exponents
Time Scale Separation Analysis
Parameter Scan
Optimization
Parameter Estimation
Sensitivities
COPASI
d([A]·V_cell)/dt = -V_cell·(k1·[A])
d([B]·V_cell)/dt = +V_cell·(k1·[A])
local parameters
display numerical value
Save Formula to Disk
Copasi Plot: Concentrations, Volumes, and Global Quantity Values
Print Save Image Save Data Zoom out Show All Hide All Close
Concentrations, Volumes, and Global Quantity Values
mol/l
s
[A] [B]
CellDesigner
File Edit Component View Database Layout Simulation Plugin Window SBW Preference Help
CellDesigner.org
Model
Compartments
Species
Reactions
Layer
base
SimpleSBML.xml *
r1
A
B
cell
Species Proteins Genes RNAs asRNAs Reactions Compartments
Edit Export
class id name speciesType compar... positio... included
PROTEIN A A cell inside Concen... 1.0
PROTEIN B B cell inside Concen... 1.0
ControlPanel SimpleSBML.xml
File Edit Data Simulation
Time span
End Time 80
Num. o... 100
Error tolerance
Exp. -6
Solver
SOSlib
COPASI
Species Parameters Change amount Parameter Sca
Id Name Compart... Quantity ... Initial
A A cell Concentra...
B B cell Concentra...
Graph Table
Concentration raw
Species Fluxes Parameters Compa
2.00
0.00
Time
80.00
V... S... C...
A
B
Unselect all
species search
search
Initialize Save As Execute Close show scatter plot

_computational BIOLOGY

RESOURCE

A community-driven global reconstruction of human metabolism

Ines Thiele[1,2,37], Neil Swainston[3,4,37], Ronan M T Fleming[1,5], Andreas Hoppe[6], Swagatika Sahoo[1], Maike K Aurich[1], Hulda Haraldsdottir[1], Monica L Mo[7], Ottar Rolfsson[1], Miranda D Stobbe[8,9], Stefan G Thorleifsson[1], Rasmus Agren[10], Christian Bölling[6], Sergio Bordel[10], Arvind K Chavali[11], Paul Dobson[12], Warwick B Dunn[3,13], Lukas Endler[14], David Hala[15], Michael Hucka[16], Duncan Hull[4], Daniel Jameson[3,4], Neema Jamshidi[7], Jon J Jonsson[5], Nick Juty[17], Sarah Keating[17], Intawat Nookaew[10], Nicolas Le Novère[17,18], Naglis Malys[3,19,20], Alexander Mazein[21], Jason A Papin[11], Nathan D Price[22], Evgeni Selkov, Sr[23], Martin I Sigurdsson[1], Evangelos Simeonidis[22,24], Nikolaus Sonnenschein[25], Kieran Smallbone[3,26], Anatoly Sorokin[21,27], Johannes H G M van Beek[28–30], Dieter Weichart[3,31], Igor Goryanin[21,32], Jens Nielsen[10], Hans V Westerhoff[3,28,33,34], Douglas B Kell[3,35], Pedro Mendes[3,4,36] & Bernhard Ø Palsson[1,7]

Multiple models of human metabolism have been reconstructed, but each represents only a subset of our knowledge. Here we describe Recon 2, a community-driven, consensus 'metabolic reconstruction', which is the most comprehensive representation of human metabolism that is applicable to computational modeling. Compared with its predecessors, the reconstruction has improved topological and functional features, including ~2× more reactions and ~1.7× more unique metabolites. Using Recon 2 we predicted changes in metabolite biomarkers for 49 inborn errors of metabolism with 77% accuracy when compared to experimental data. Mapping metabolomic data and drug information onto Recon 2 demonstrates its potential for integrating and analyzing diverse data types. Using protein expression data, we automatically generated a compendium of 65 cell type–specific models, providing a basis for manual curation or investigation of cell-specific metabolic properties. Recon 2 will facilitate many future biomedical studies and is freely available at http://humanmetabolism.org/.

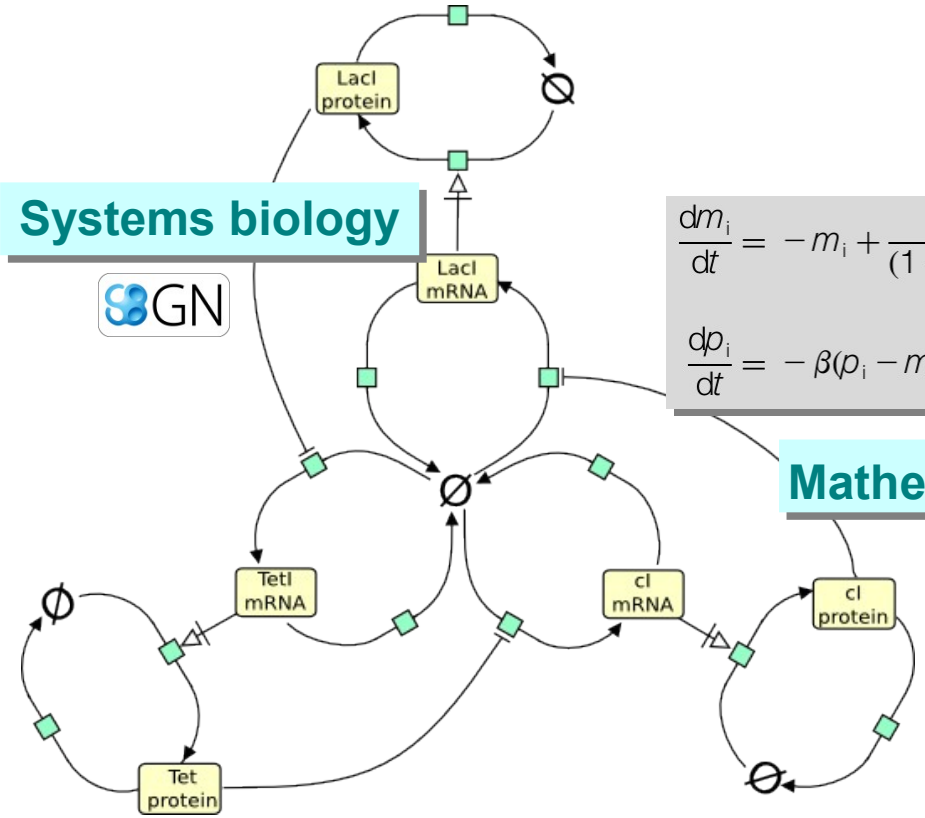
An understanding of metabolism is fundamental to comprehending the phenotypic behavior of all living organisms, including humans, where metabolism is integral to health and is involved in much of human disease. High quality, genome-scale 'metabolic reconstructions' are at the heart of bottom-up systems biology analyses and represent the entire network of metabolic reactions that a given organism is known to exhibit[1]. The metabolic-network reconstruction procedure is now well-established[2] and has been applied to a growing number of model organisms[3]. Metabolic reconstructions allow for the conversion of biological knowledge into a mathematical format and the subsequent computation of physiological states[1,4,5] to address a variety of scientific and applied questions[3,6]. Reconstructions enable network-wide mechanistic investigations of the genotype-phenotype relationship. A high-quality reconstruction of the metabolic network is thus

A not so simple SBML file (Recon2)

- 8 compartments
- 5 063 metabolites
- 2 194 proteins
- 7 440 reactions

MODEL1109130000

BioModels Database – http://www.ebi.ac.uk/biomodels



Systems biology

SBGN

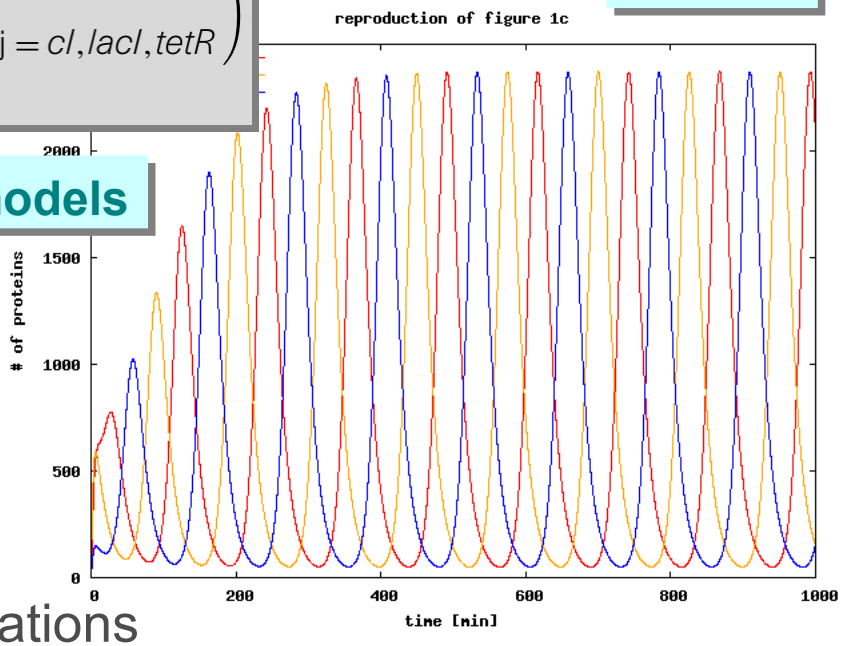
$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \begin{pmatrix} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{pmatrix}$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i)$$

Mathematical models

SBML

curated



>140 000 models, >10 millions mathematical relations

Deposition advised by >300 journals, database > 600 citations

1.5 million page requests per year

Submission in SBML and CellML; Export in SBML, CellML, XPP, SciLab, BioPAX, Octave, PDF, VCML, SBGN

Babraham Institute

BioModels Database

Search
Advanced

BioModels Home | Models | Submit | Support | About BioModels | Contact us

BioModels Database serves as a reliable repository of computational models of biological processes. It hosts models described in peer-reviewed scientific literature and models generated automatically from pathway resources (Path2Models). A large number of models collected from literature are manually curated and semantically enriched with cross-references from external data resources. The resource allows scientific community to store, search and retrieve mathematical models of their interest. In addition, features such as generation of sub-models, online simulation, conversion of models into different representational formats, and programmatic access via web services, are provided.

Models published in the literature

- Browse curated models
- Browse curated m… (Browse curated models)
- Browse curated models using Taxonomy
- Browse non-curated models

Browse the models in the curated branch. These models have been thoroughly curated, and model elements have been annotated with terms from controlled vocabularies as well as links to relevant data resources.

Path2Models

Submit a model

Links

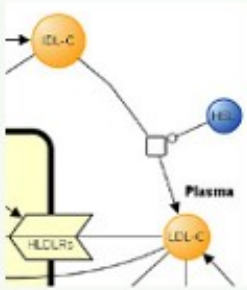
- Main instance at EMBL-EBI, UK
- Mirror at Caltech, USA
- Project on SourceForge
- Web Services
- Download archived models

http://www.ebi.ac.uk/biomodels

Model of the month

October, 2013

Why LDL-C rises with age? The model presented here demonstrates that mechanistic changes to cholesterol absorption and LDL-C removal from the plasma with age, can result in a significant rise in LDL-C.



Please read more...

News

Follow us on Twitter

8 October 2013 **SPARQL endpoint available**

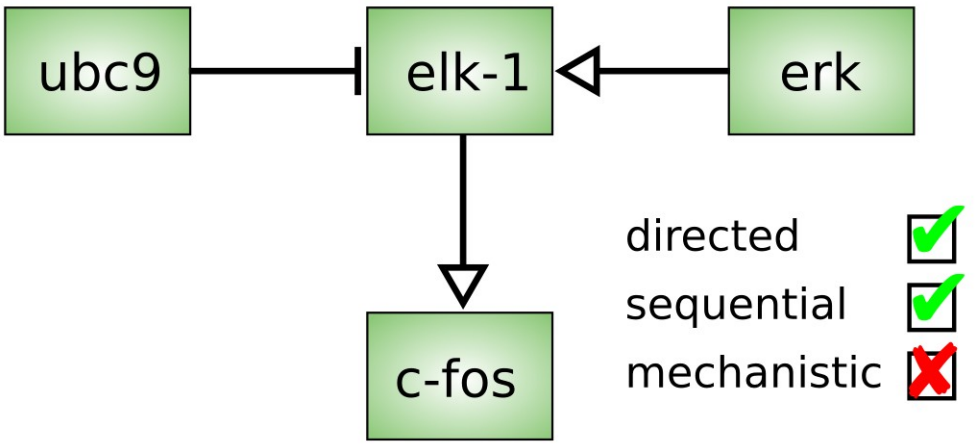
The EBI launched a RDF platform. This provides access to various bioinformatics resources (including BioModels Database) using Semantic Web technologies (SPARQL endpoints and RDF dumps are available). Read more about this news.

18 June 2013 **25th release**

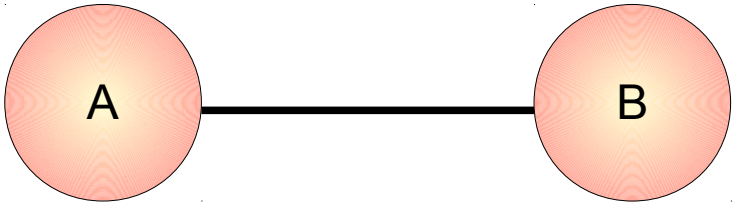
We are pleased to announce the twenty-fifth release of BioModels Database! The resource now publicly provides 143013 models. Numerous new models have been made available, several have been updated and various new

Activity Flows

- Logical modelling
- Signalling pathways, gene regulatory networks

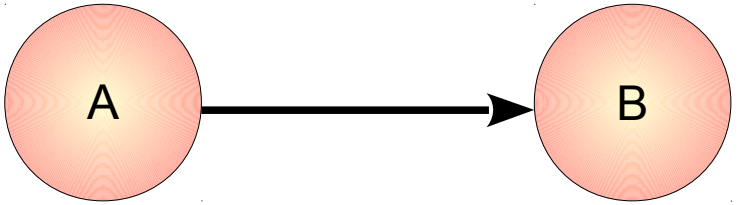


Undirected, directed, signed



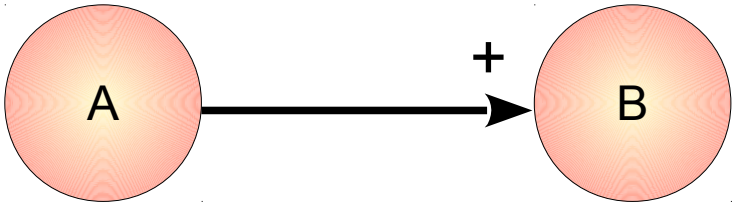
Undirected

"A interacts with B"



directed

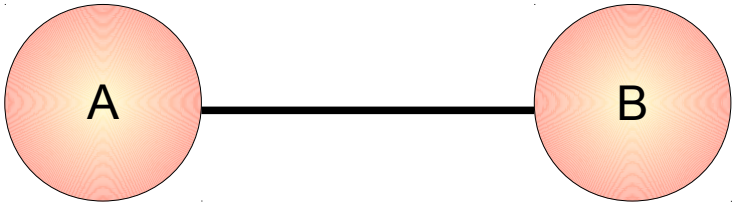
"A influences B"

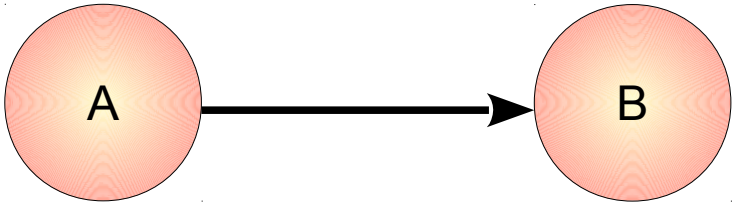


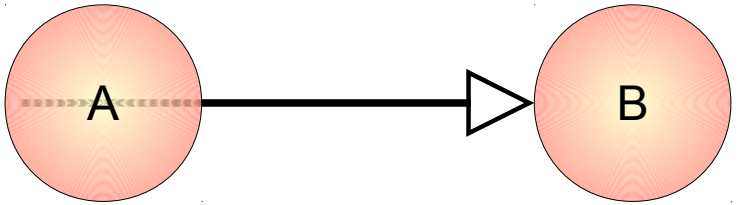
Signed

A influences positively B

Undirected, directed, signed

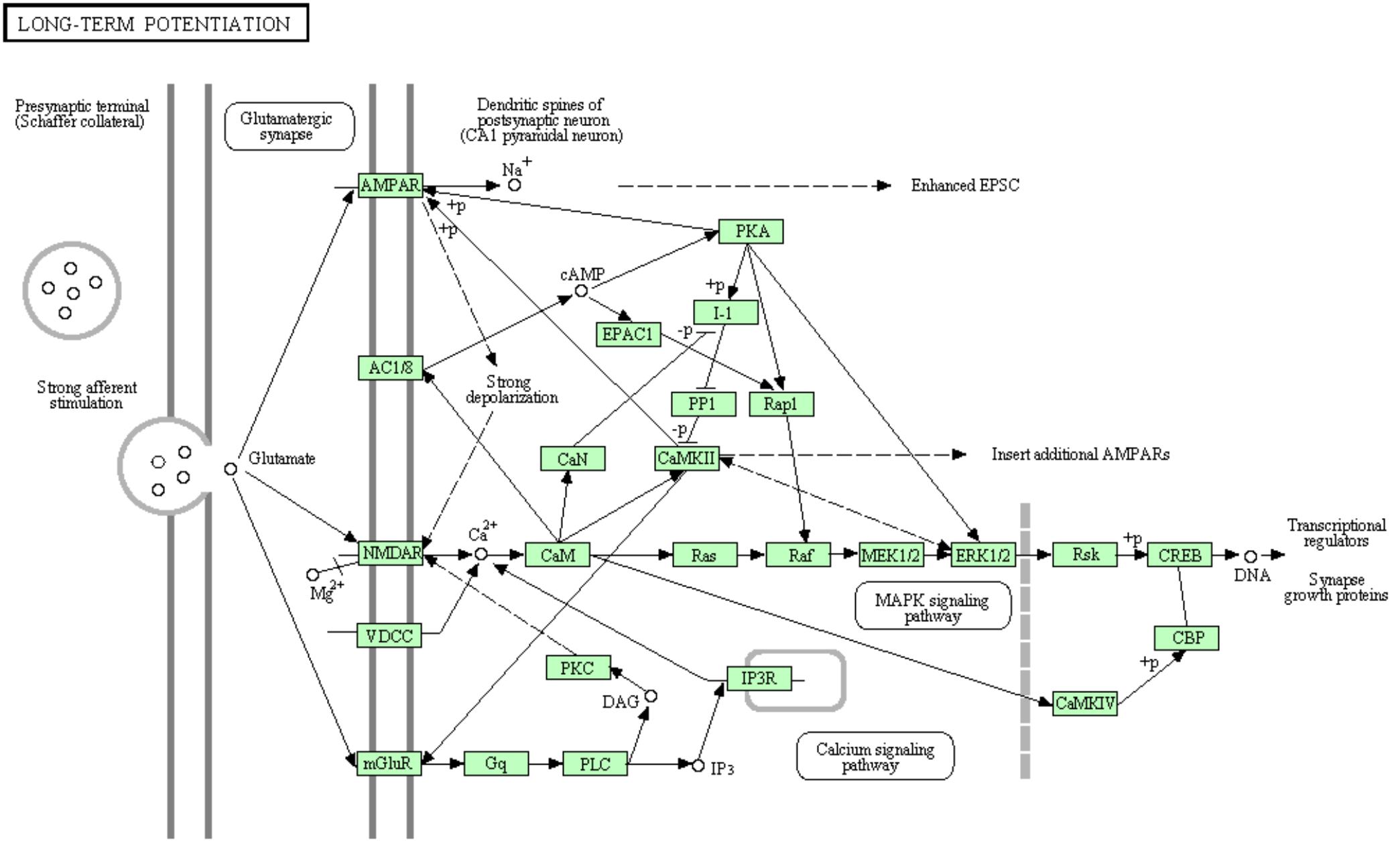






A signed interaction network
is equivalent to an activity flow

http://www.genome.jp/kegg/pathway.html



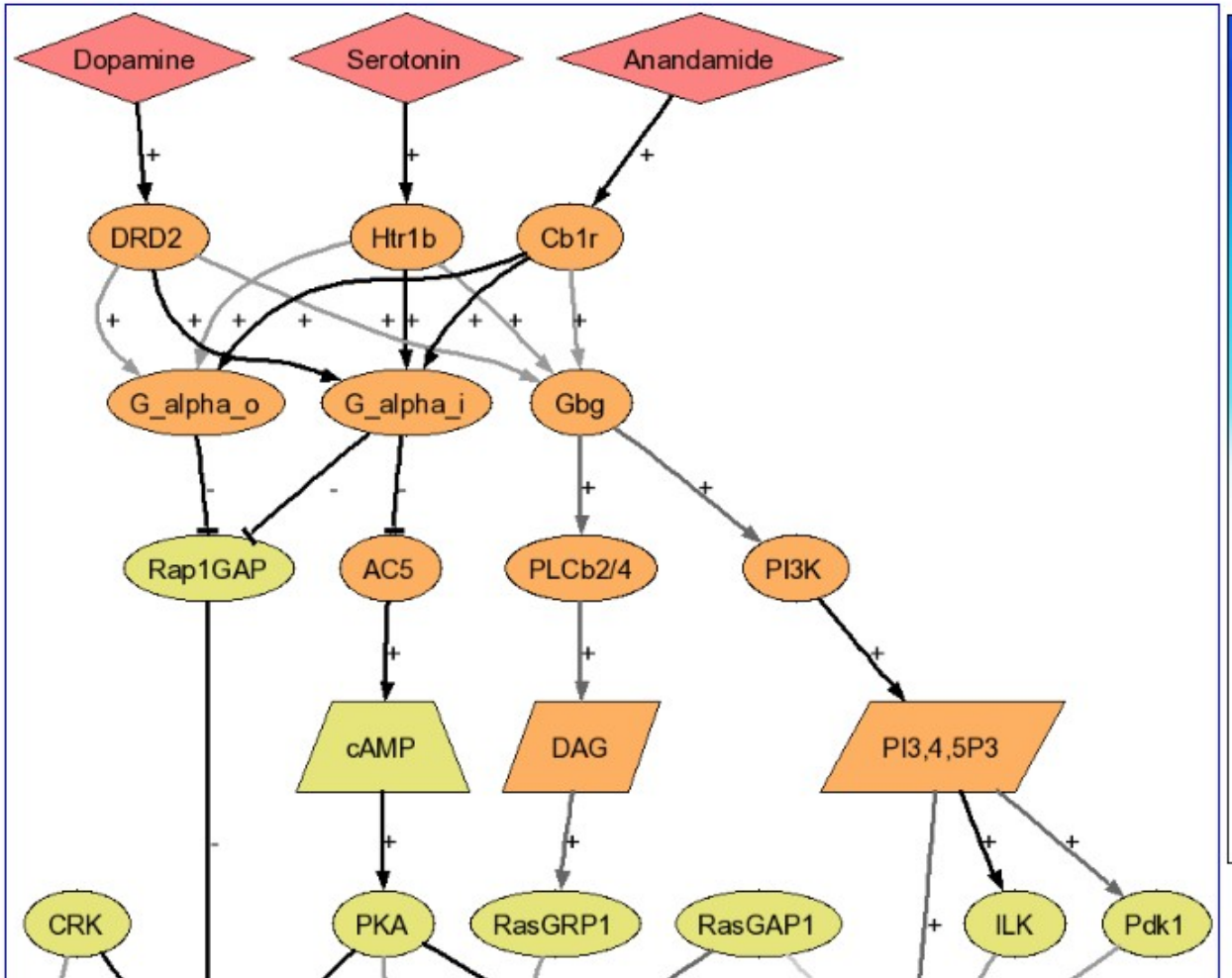
http://stke.sciencemag.org/cm/

Pathway Information

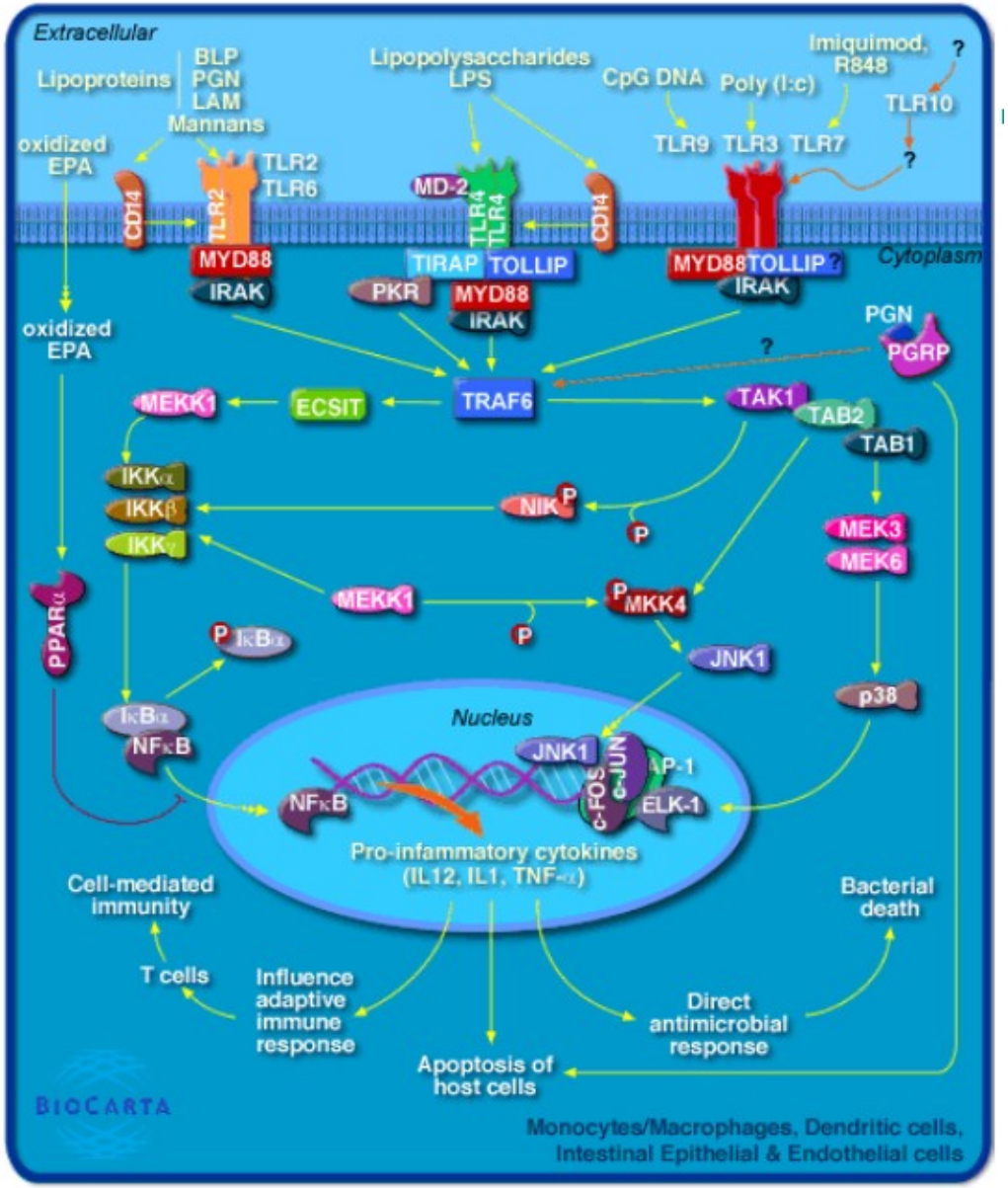
Neuro2A Differentiation by G alpha i/o Pathway

Authorities: Avi Ma'ayan, et al.
Organism: vertebrates : mammals : Rodentia : mice
Tissue Cell: nervous system : central nervous system : neurons
Citations: Literature
Scope: Specific
Canonical Pathways: G alpha i Pathway

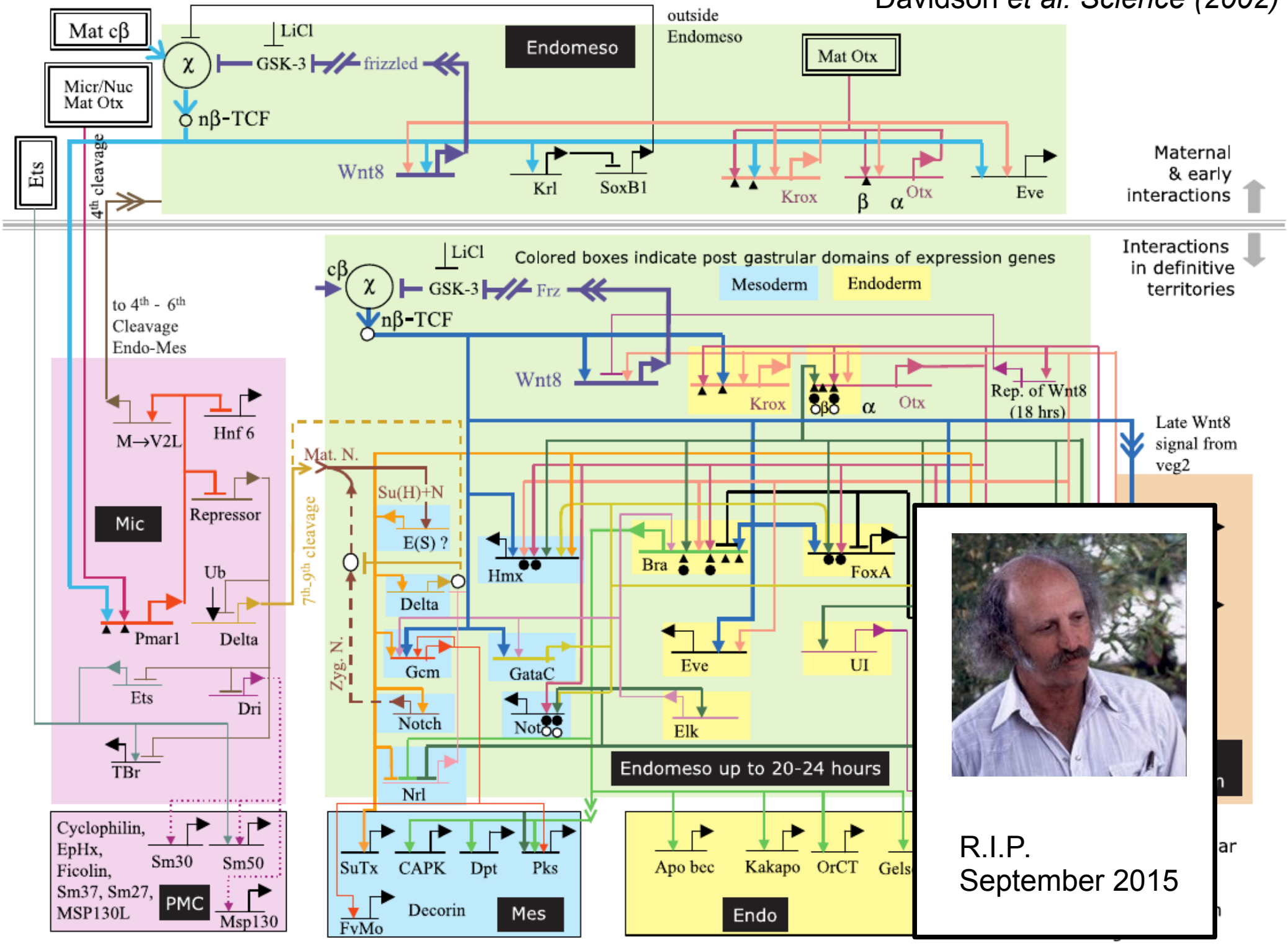
Pathway
View Connections Map Legend
Get SVG Pathway Image



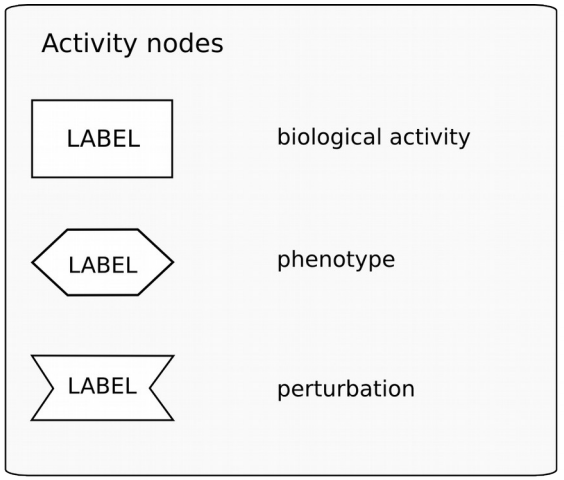
http://cgap.nci.nih.gov/Pathways/BioCarta.org

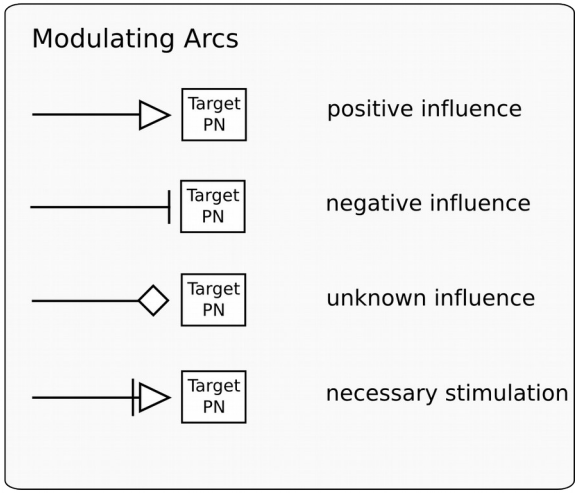


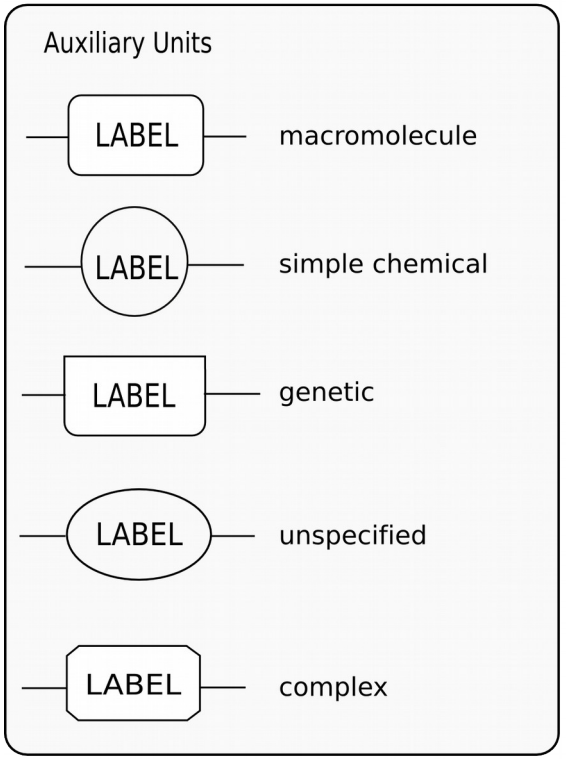
Davidson *et al. Science (2002)*

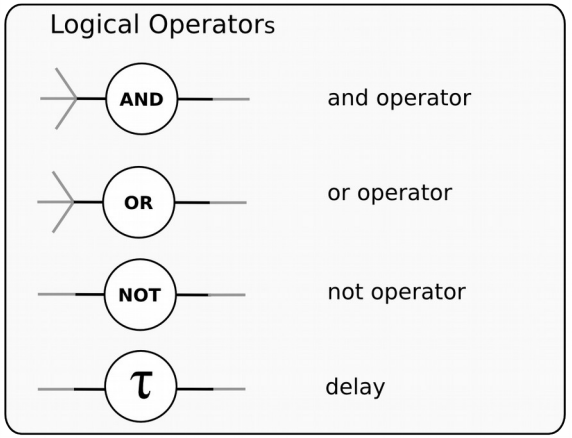


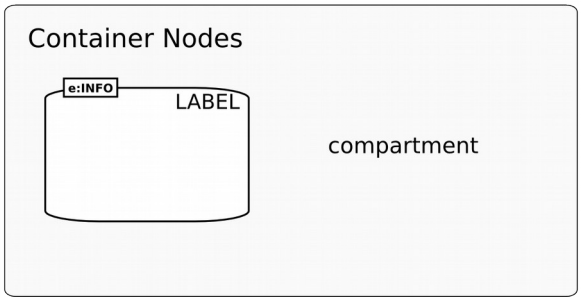
SBGN Activity Flows L1 reference card



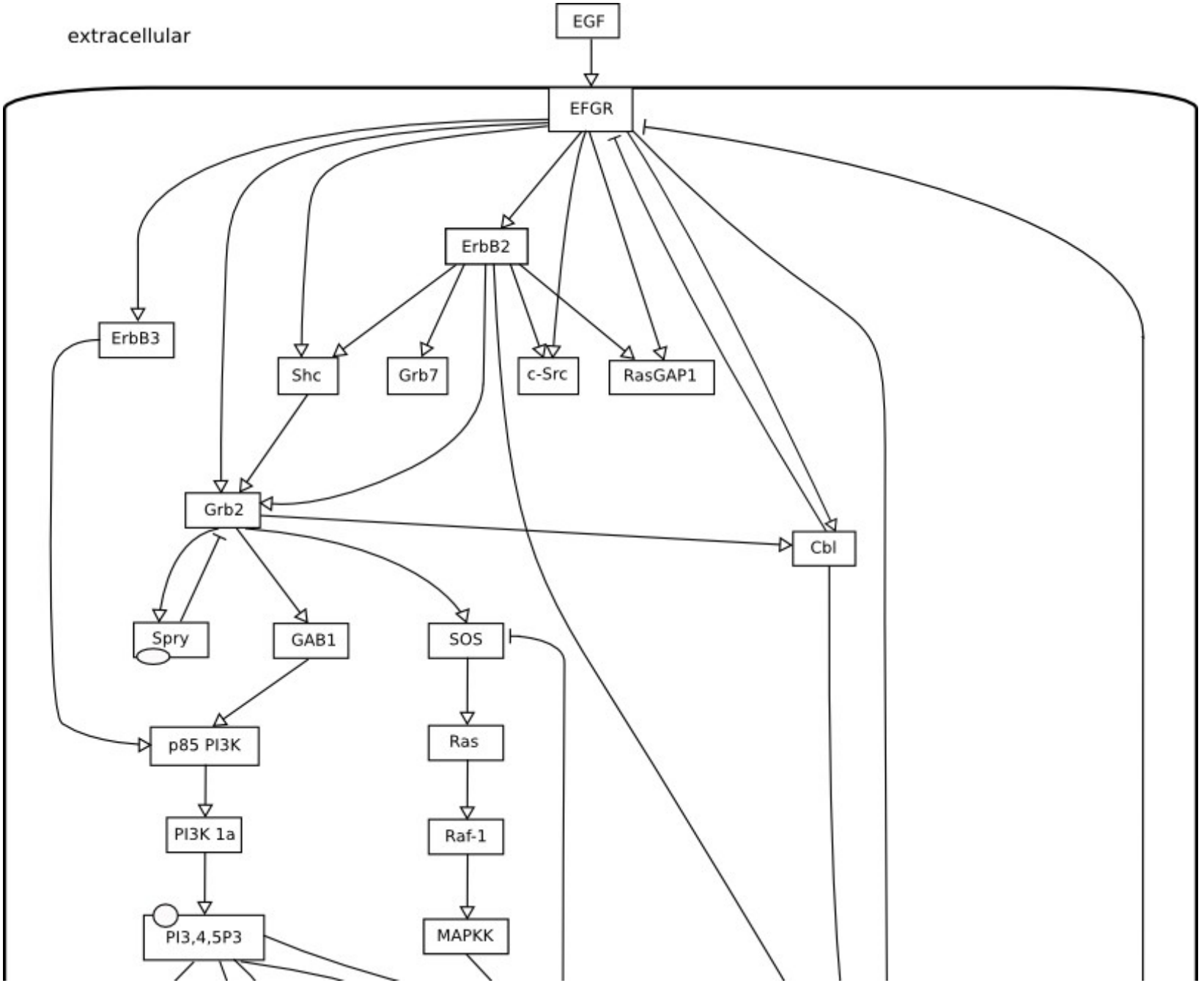




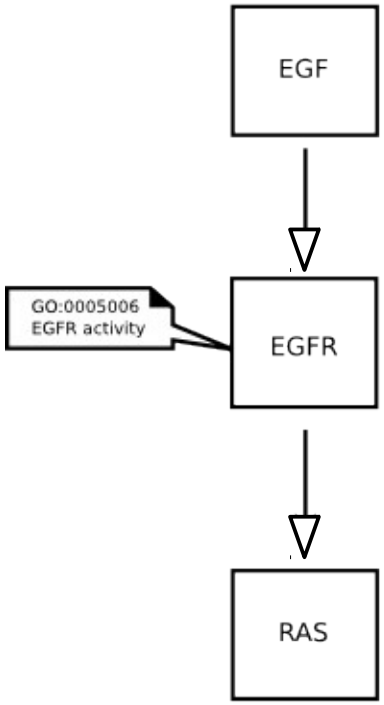




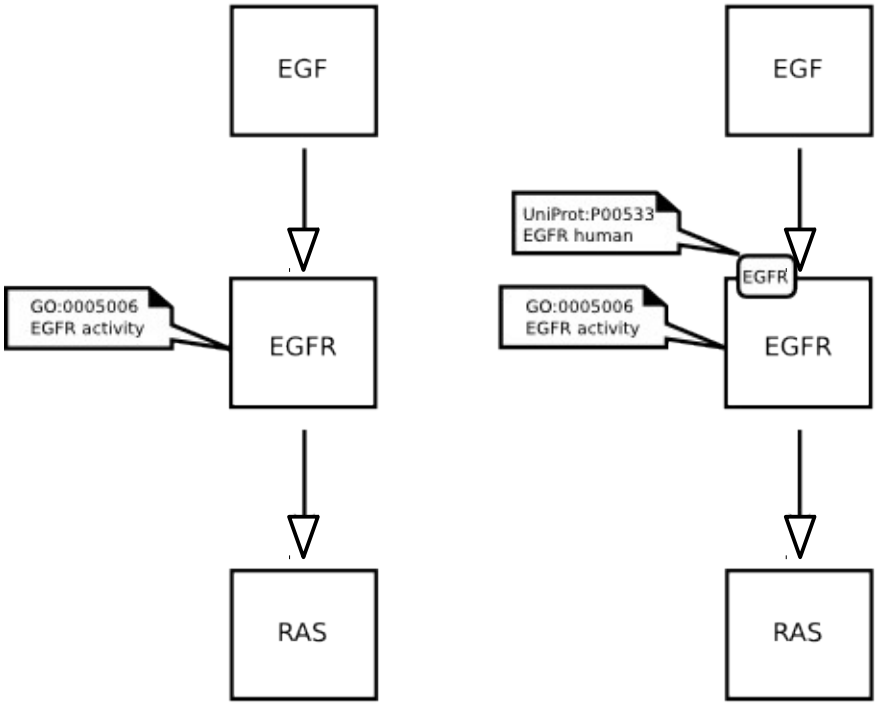
Example of Activity Flow map



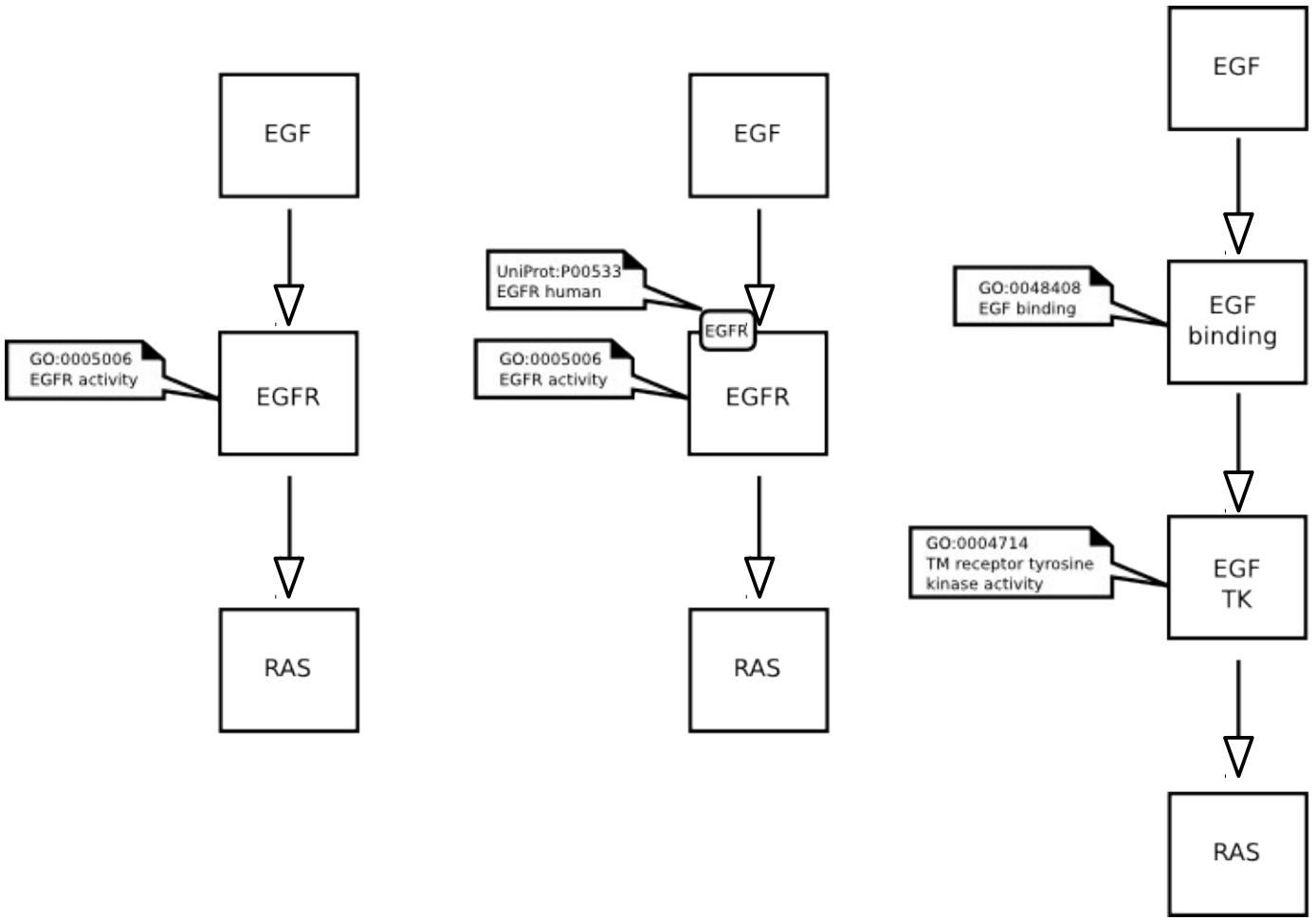
TIMTOWTDI
(There Is More Than One Way To Do It)



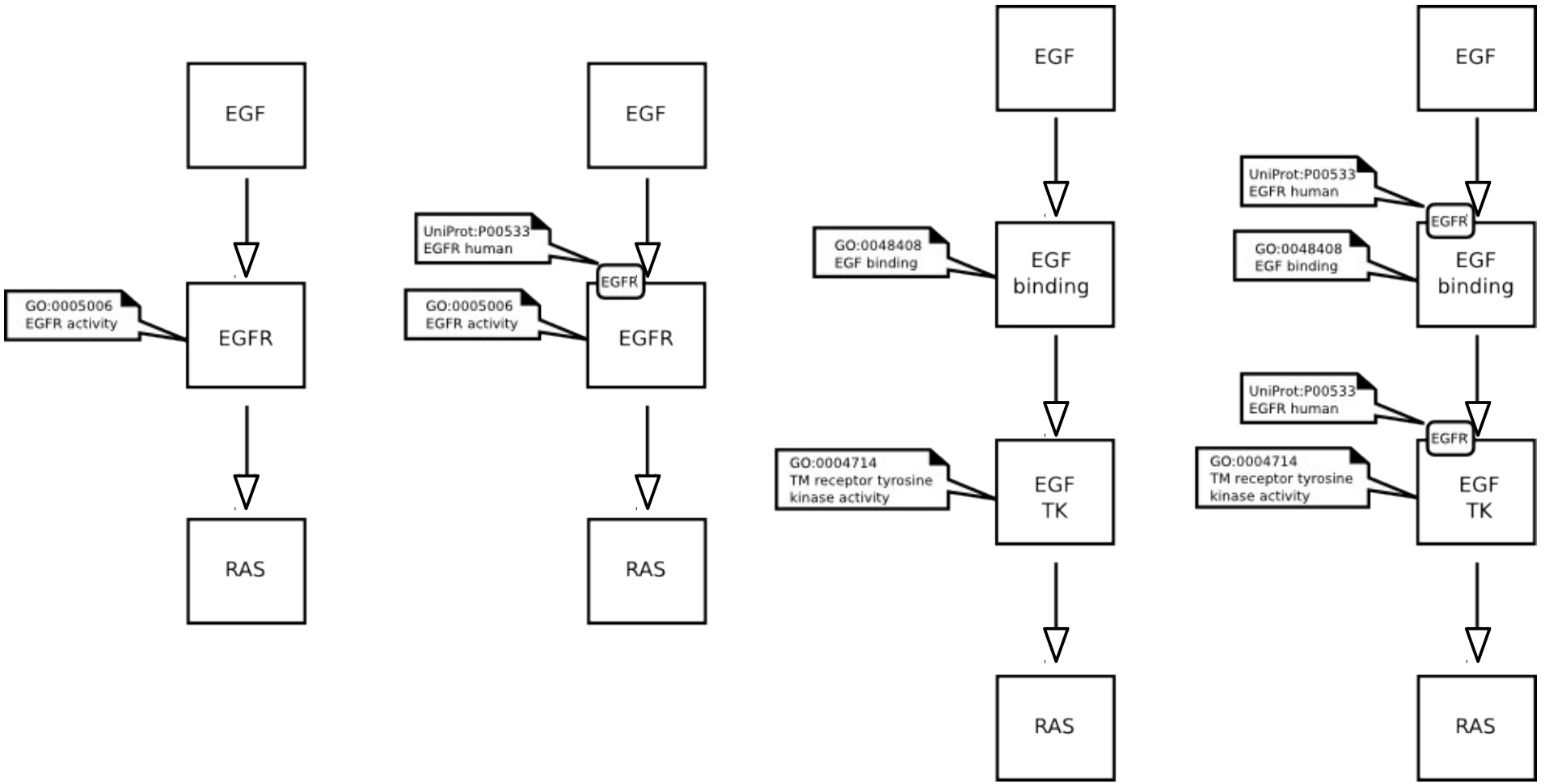
TIMTOWTDI


EGF
GO:0005006
EGFR activity
EGFR
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0005006
EGFR activity
EGFR
RAS

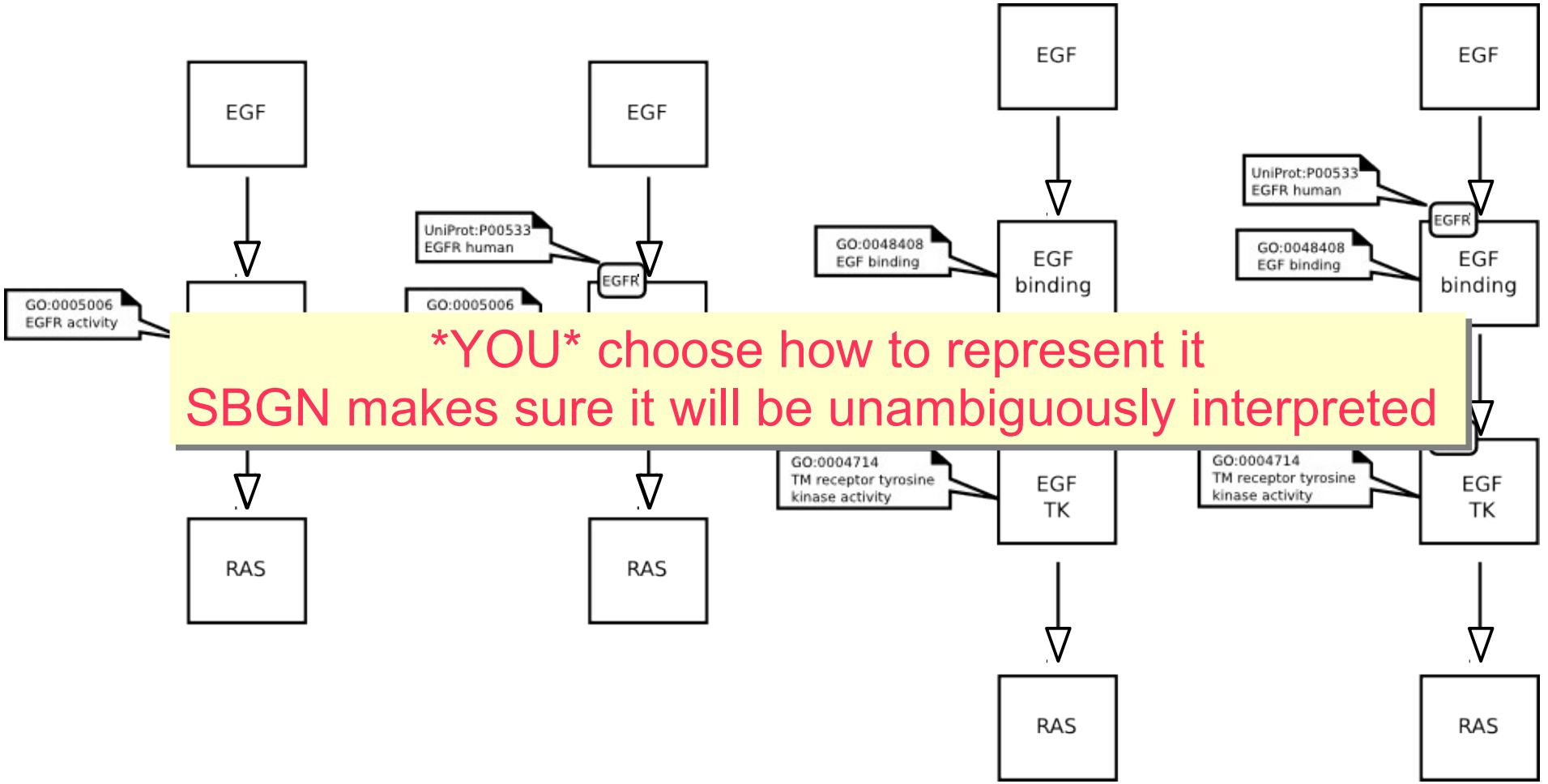
TIMTOWTDI


EGF
GO:0005006
EGFR activity
EGFR
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0005006
EGFR activity
EGFR
RAS
EGF
GO:0048408
EGF binding
EGF
binding
GO:0004714
TM receptor tyrosine
kinase activity
EGF
TK
RAS

TIMTOWTDI

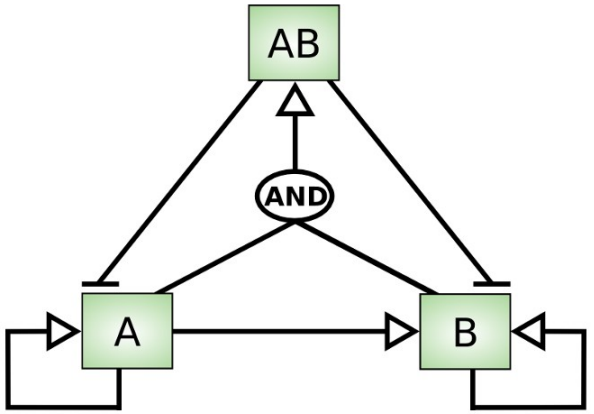

EGF
GO:0005006
EGFR activity
EGFR
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0005006
EGFR activity
EGFR
RAS
EGF
GO:0048408
EGF binding
EGF
binding
GO:0004714
TM receptor tyrosine
kinase activity
EGF
TK
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0048408
EGF binding
EGF
binding
UniProt:P00533
EGFR human
EGFR
GO:0004714
TM receptor tyrosine
kinase activity
EGF
TK
RAS

TIMTOWTDI

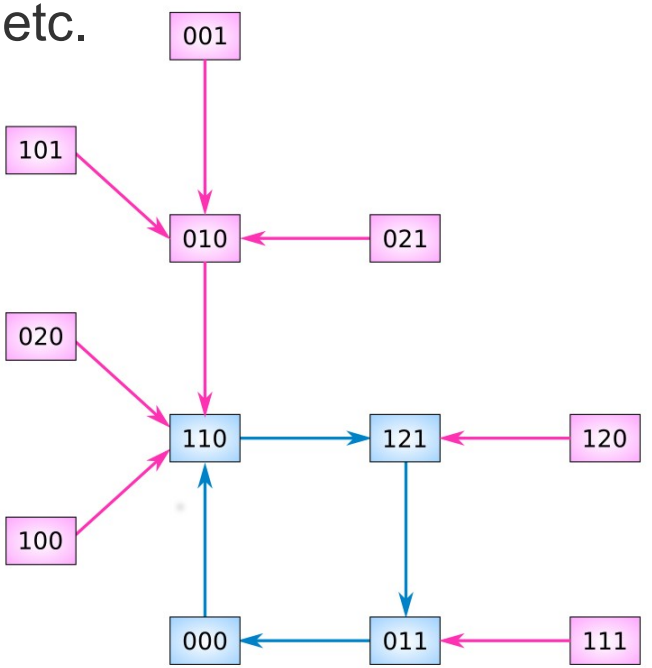

EGF
EGF
EGF
EGF
UniProt:P00533
EGFR human
UniProt:P00533
EGFR human
EGFR
EGFR
GO:0005006
EGFR activity
GO:0005006
GO:0048408
EGF binding
GO:0048408
EGF binding
EGF
binding
EGF
binding
YOU choose how to represent it
SBGN makes sure it will be unambiguously interpreted
GO:0004714
TM receptor tyrosine
kinase activity
GO:0004714
TM receptor tyrosine
kinase activity
EGF
TK
EGF
TK
RAS
RAS
RAS
RAS

Logic models

- Variables can take a discrete number of values, at least 2
- Transitions of output are expressed as logical combinations of input values
- Simulations can be:
  synchronous: all the nodes are updated at once
  asynchronous: nodes are updated one after the other
- One can add delays, inputs etc.



Influence diagram

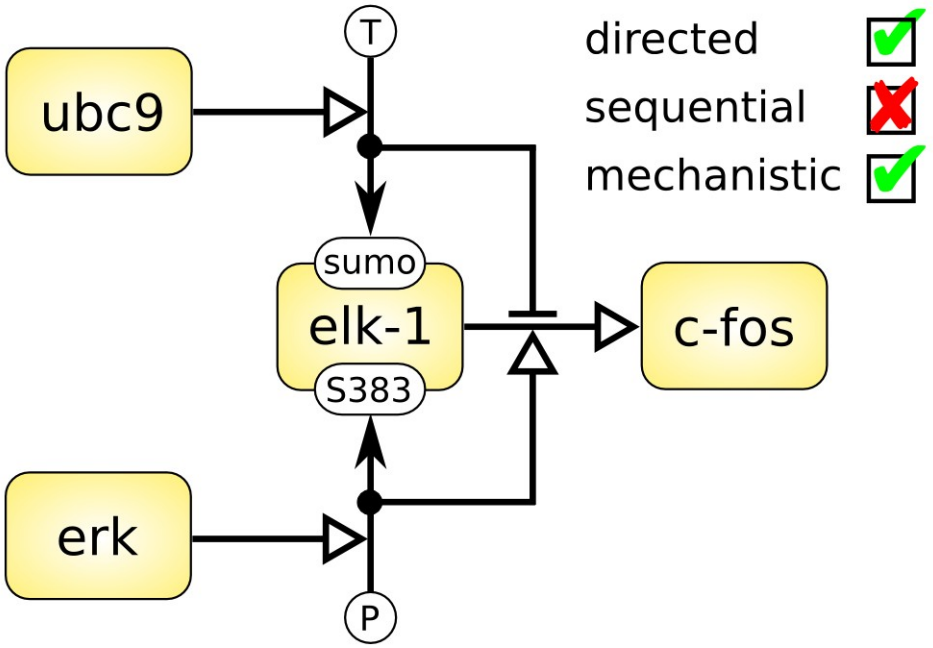


state diagram

SBML Qual

Babraham Institute

Entity Relationships



- Rule-based modelling
- Molecular Biology
- “Open world”
- Independent rules: no explosion

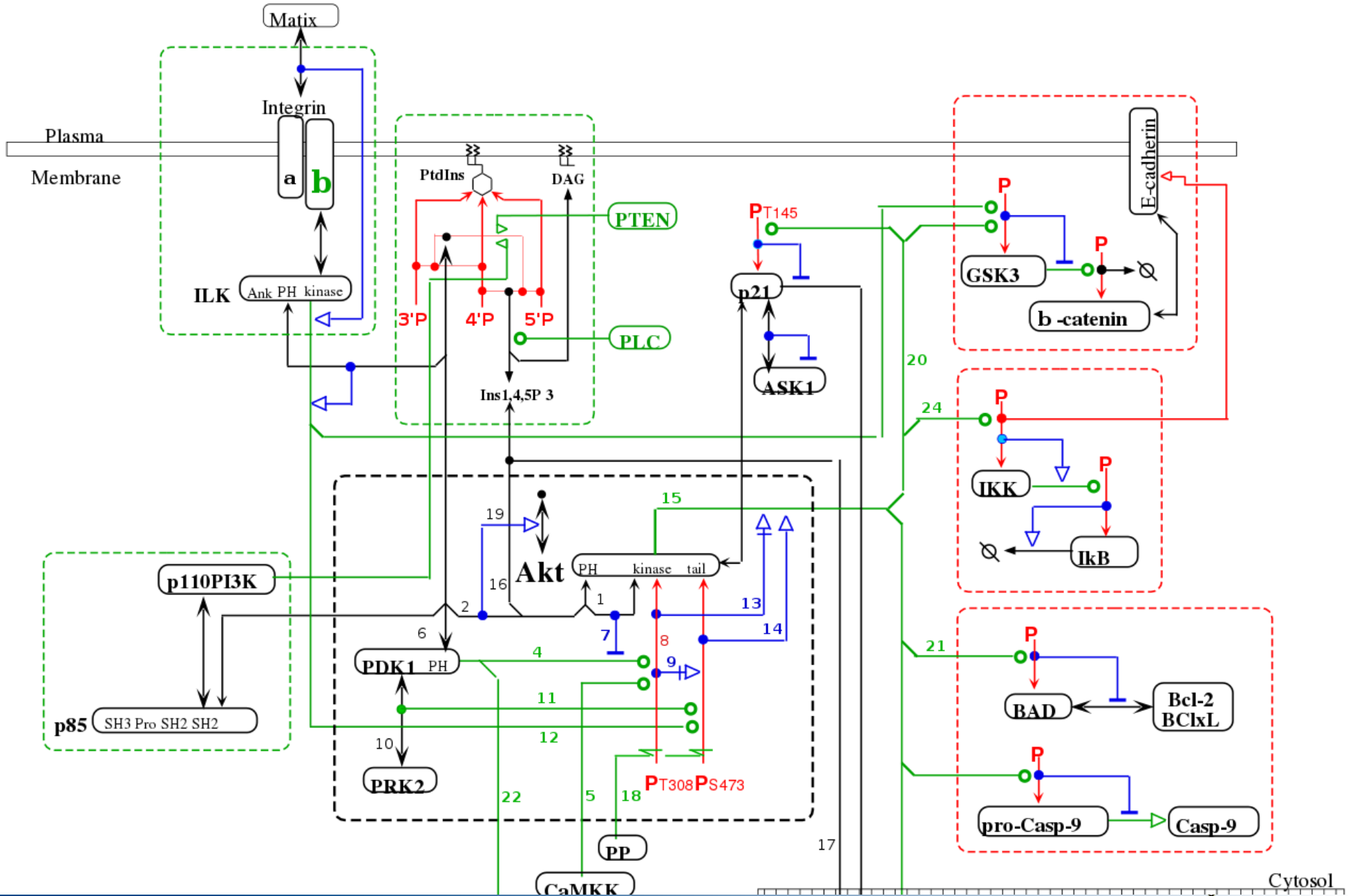
http://discover.nci.nih.gov/mim/

Map description

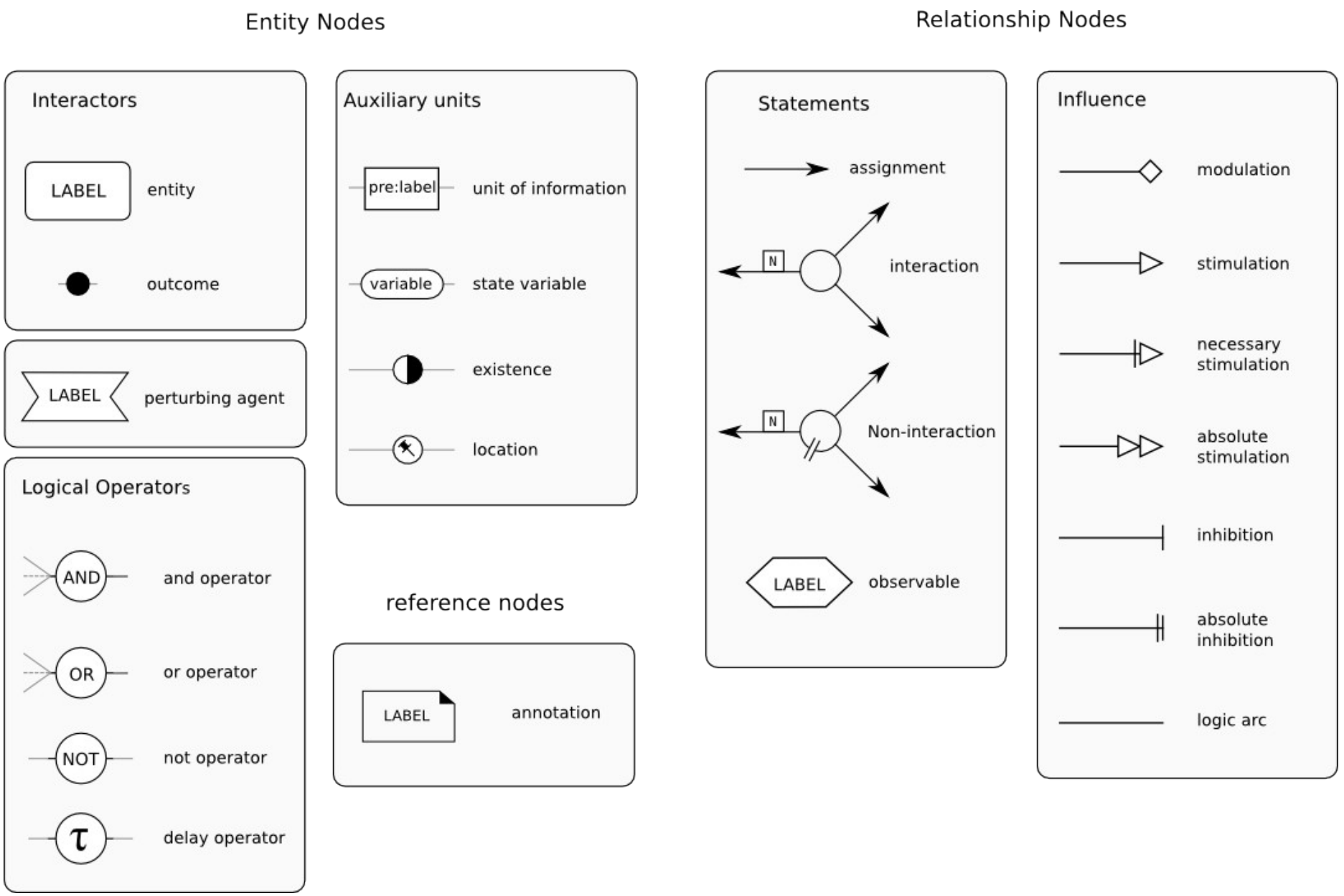
(see System Requirements)

AKT

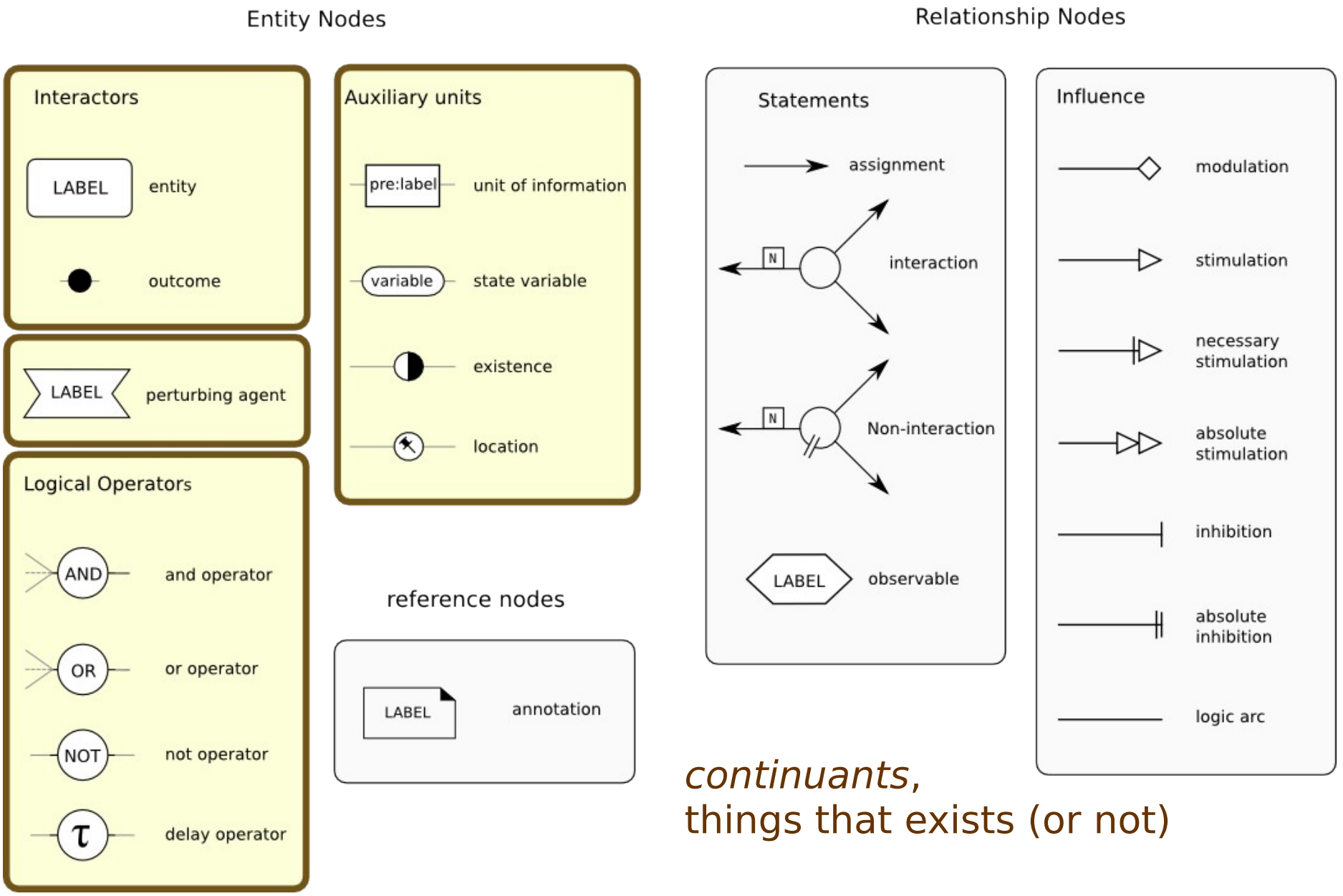
Click here to return to original map size



SBGN Entity Relationships L1 reference card

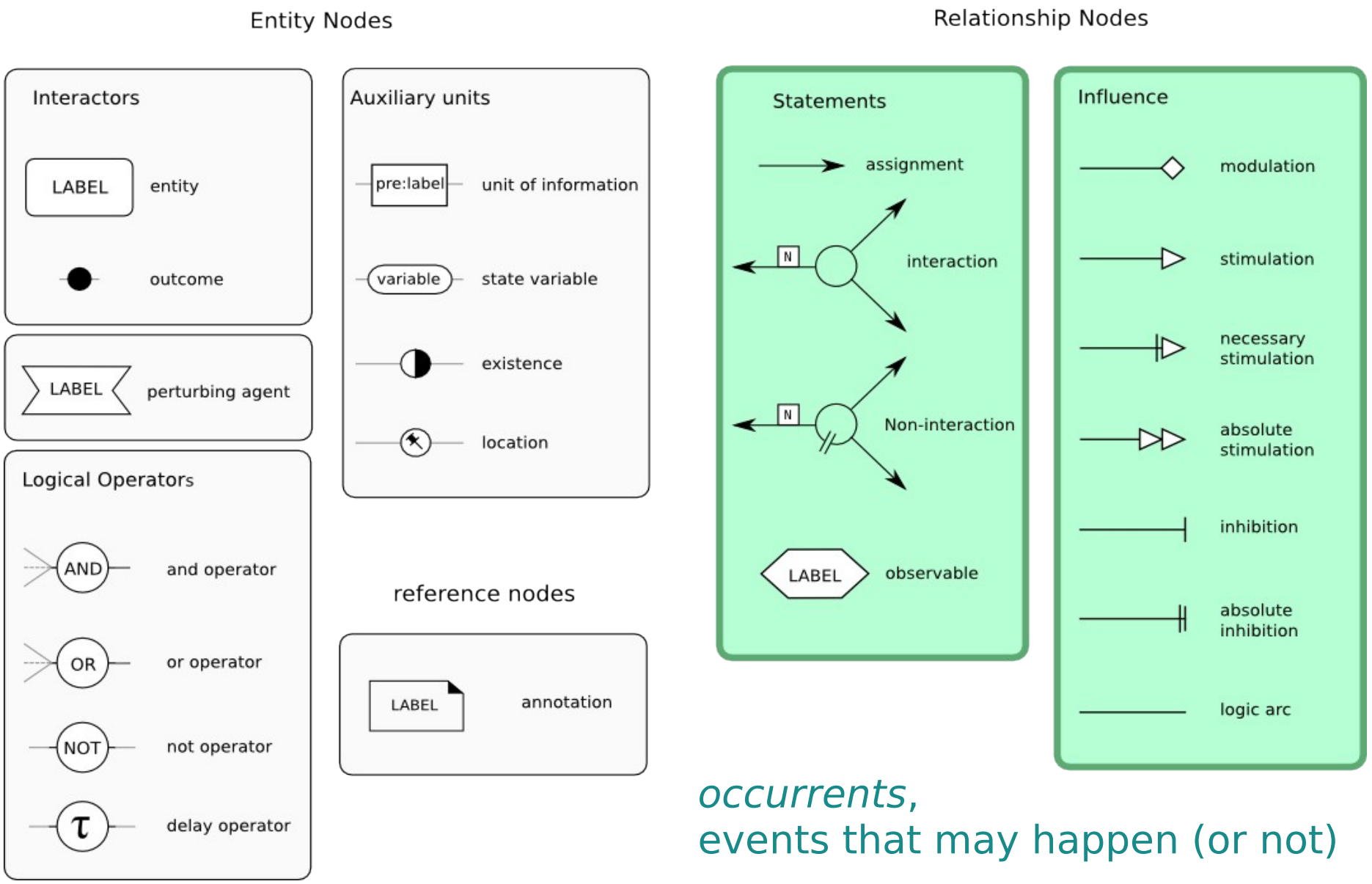


SBGN Entity Relationships L1 reference card



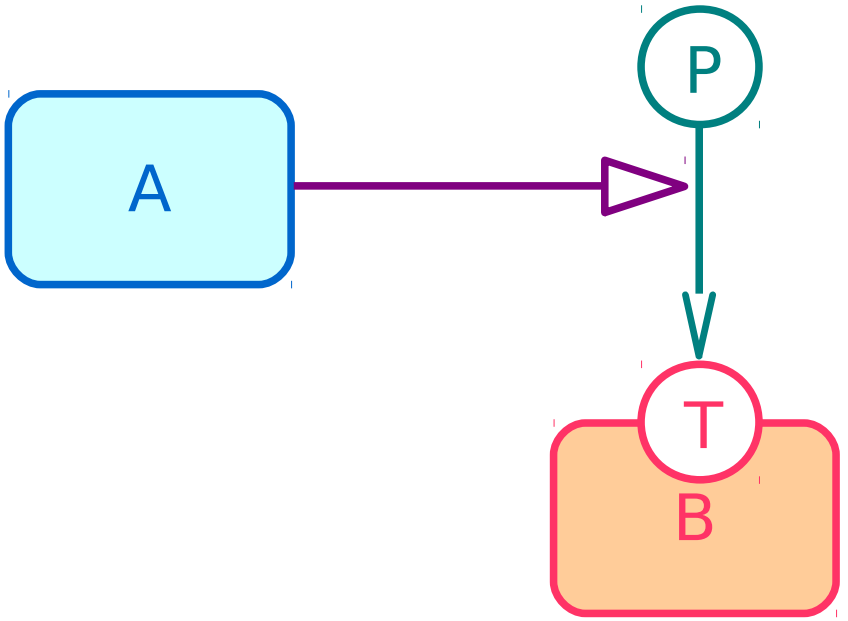
continuants,
things that exists (or not)

SBGN Entity Relationships L1 reference card

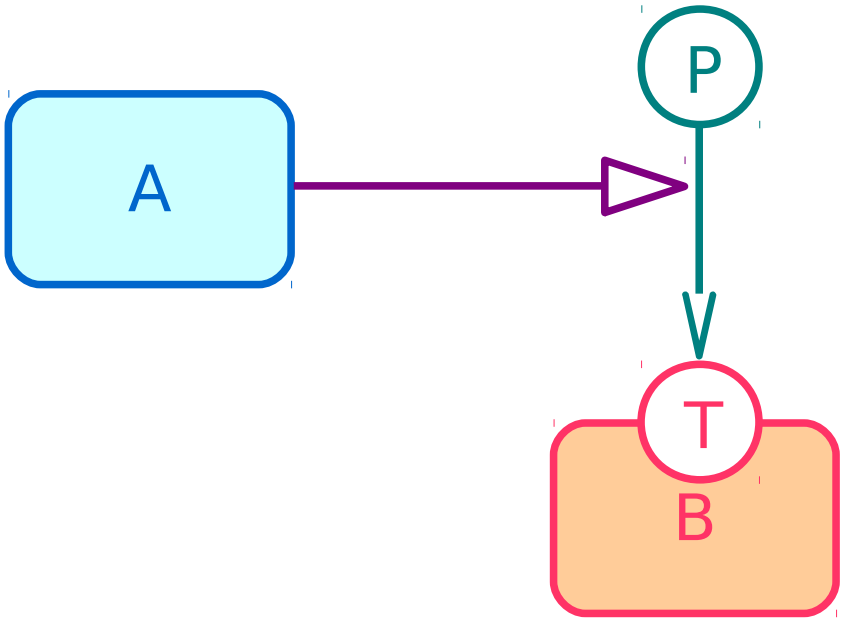


occurrents,
events that may happen (or not)

Entity Relationships can be viewed as rules

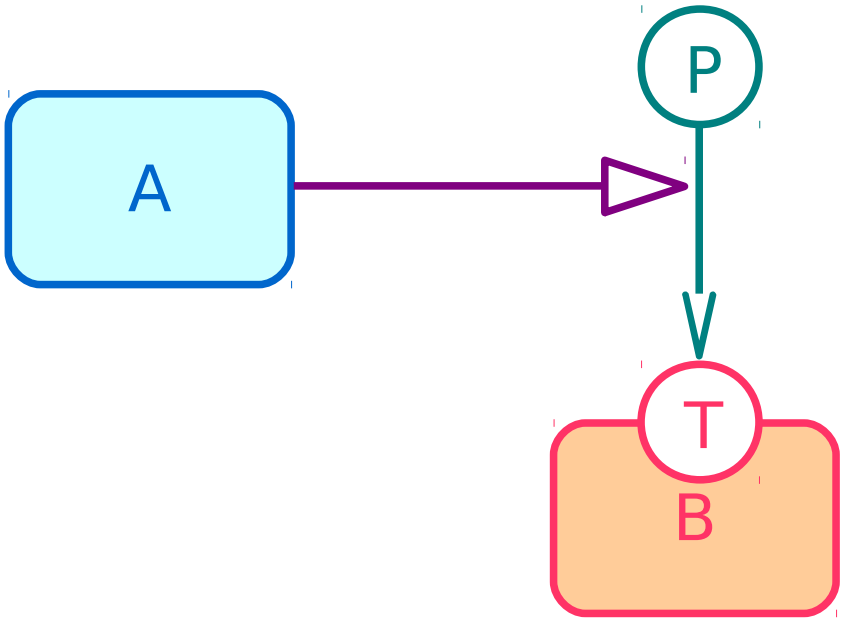


Entity Relationships can be viewed as rules



If A exists, the assignment of the value P to the state variable T of B is increased

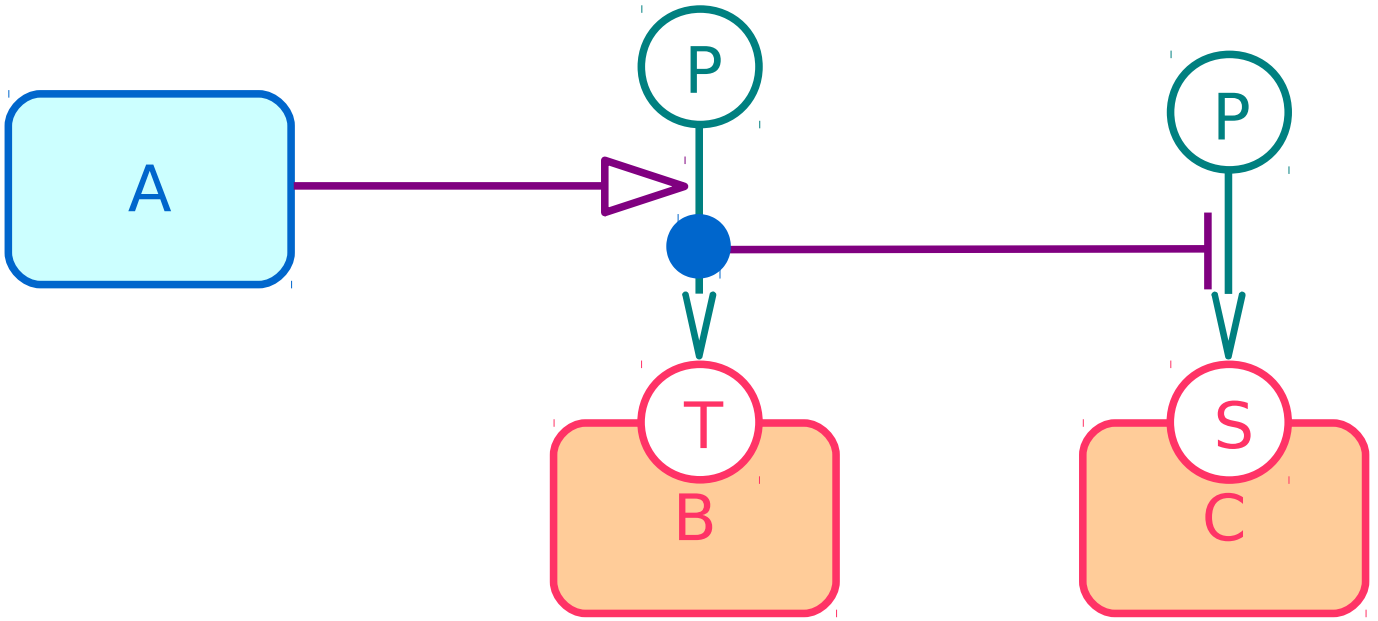
Entity Relationships can be viewed as rules



If A exists, the assignment of the value P to the state variable T of B is increased

(A stimulates the phosphorylation of B on the threonine)

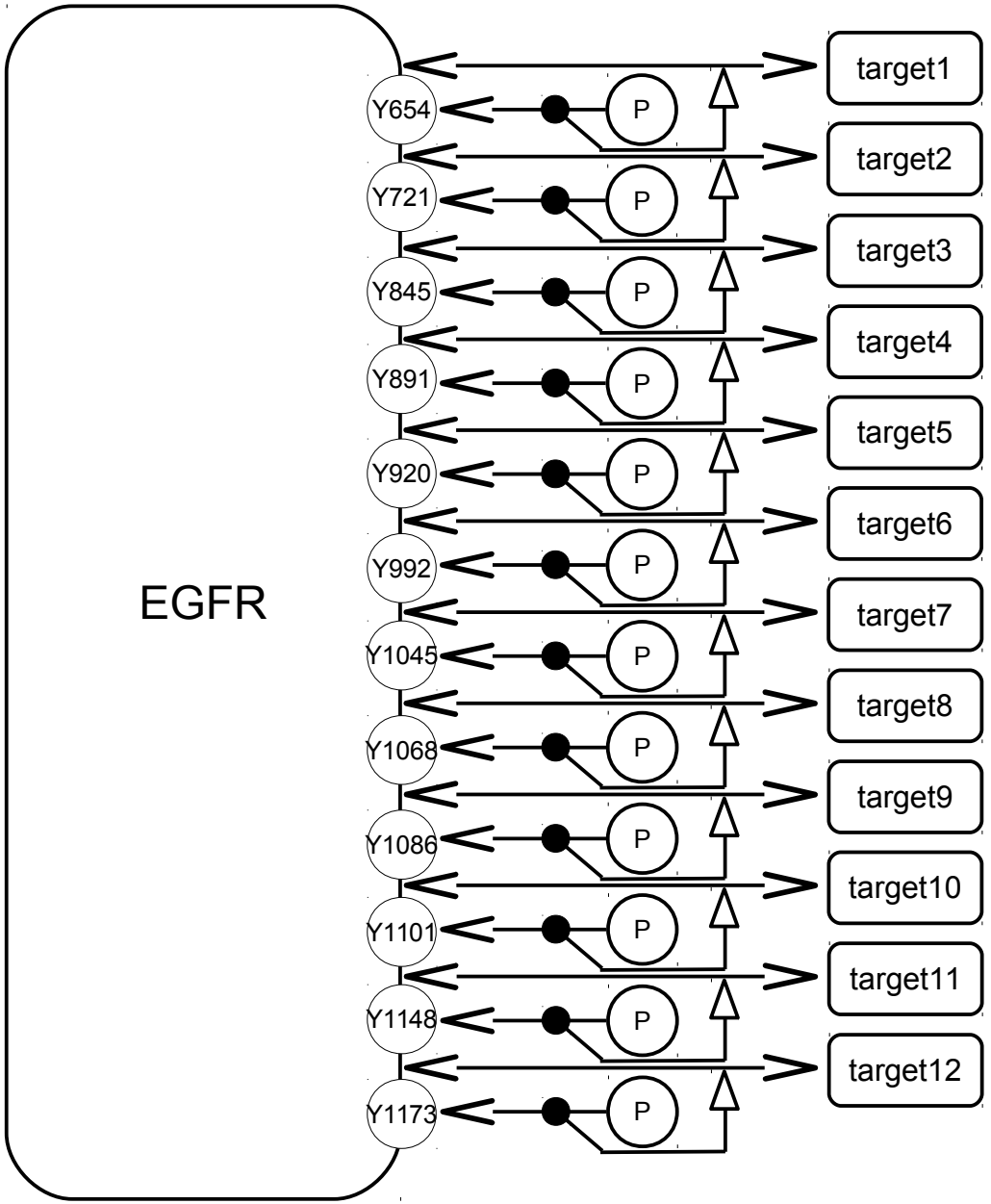
Entity Relationships can be viewed as rules



If A exists, the assignment of the value P to the state variable T of B is increased

If P is assigned to the state variable T of B, the assignment of the value P to the state variable S of B is decreased

Multistate and combinatorial explosion



Process Descriptions:
"once a state variable value, always a state variable value"

2^{12} = 4096 states
(i.e. EPN glyphs) for EGFR
and 4096 complexes between
EGFR and targets

Rule-based modelling

- Kappa http://www.kappalanguage.org/

  Kappa: A(Site1~u),B(Site1) -> A(Site1~u!1),B(Site1!1)

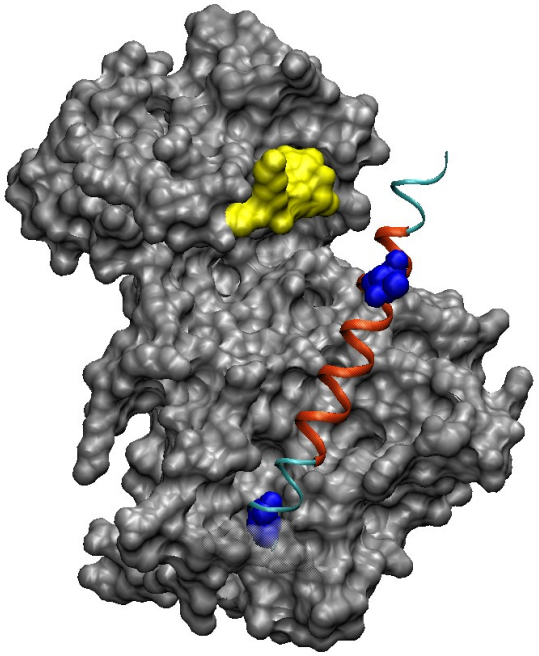
  English: "Unphosphorylated Site1 of A binds to Site1 of B"

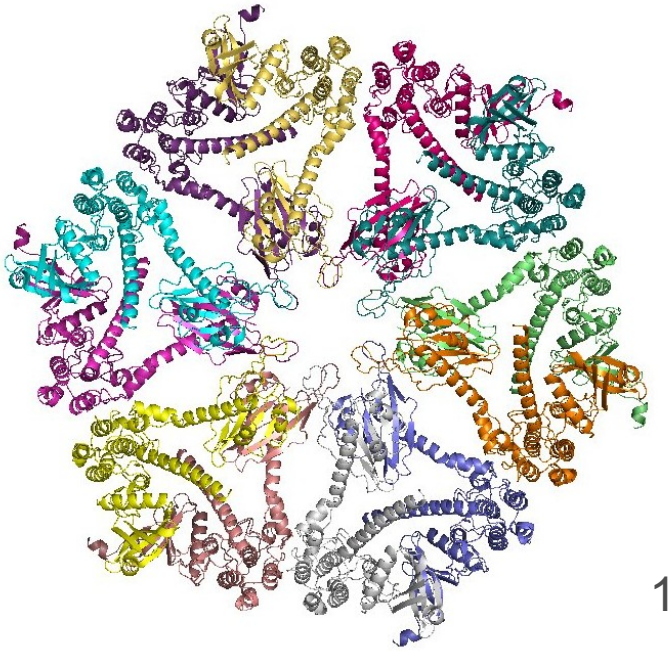
- BioNetGen http://bionetgen.org/

  EGF(R) + EGFR(L) <-> EGF(R!1).EGFR(L!1) kp1, km1

  English: “unbound EGF receptor site binds to unbound receptor ligand site”

Multi-state nature of CaMKII and rule-based modelling



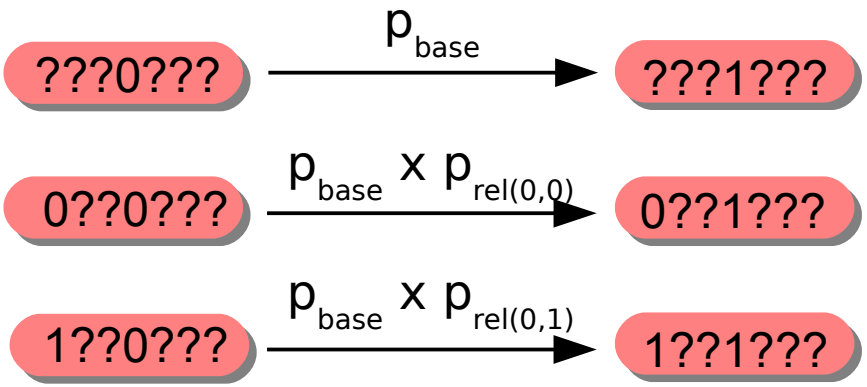


12 monomers

5 binary state variables

Open/Closed
T286 P/nonP
T306 P/nonP
CaM on low affinity
CaM on high affinity

2^{60} states = 1 billions of billions



Flags '?' do not affect the reaction

4 species are necessary and not 128

2 reactions are necessary and not 64

Simulation with StochSim

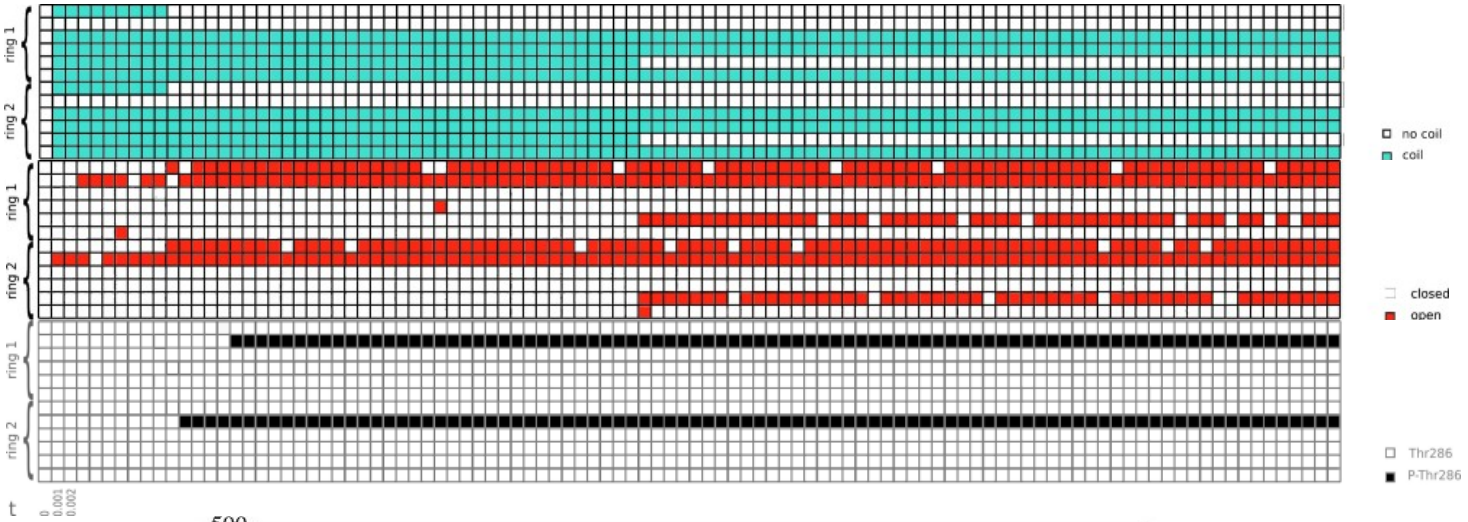
time

single oligomer

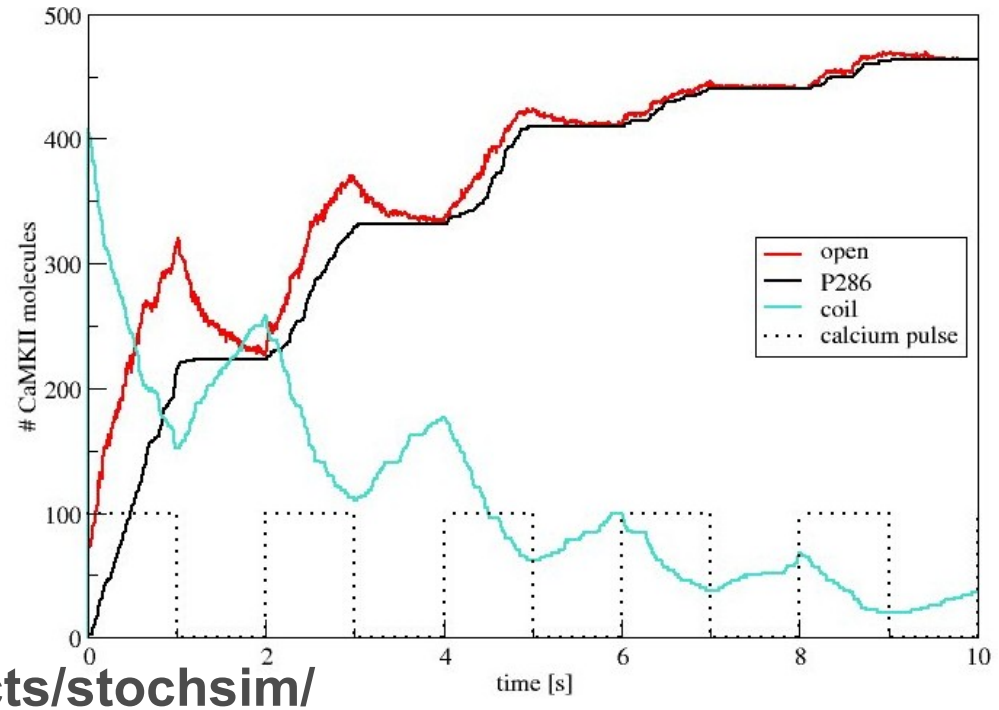
dimer coil

open

T286P

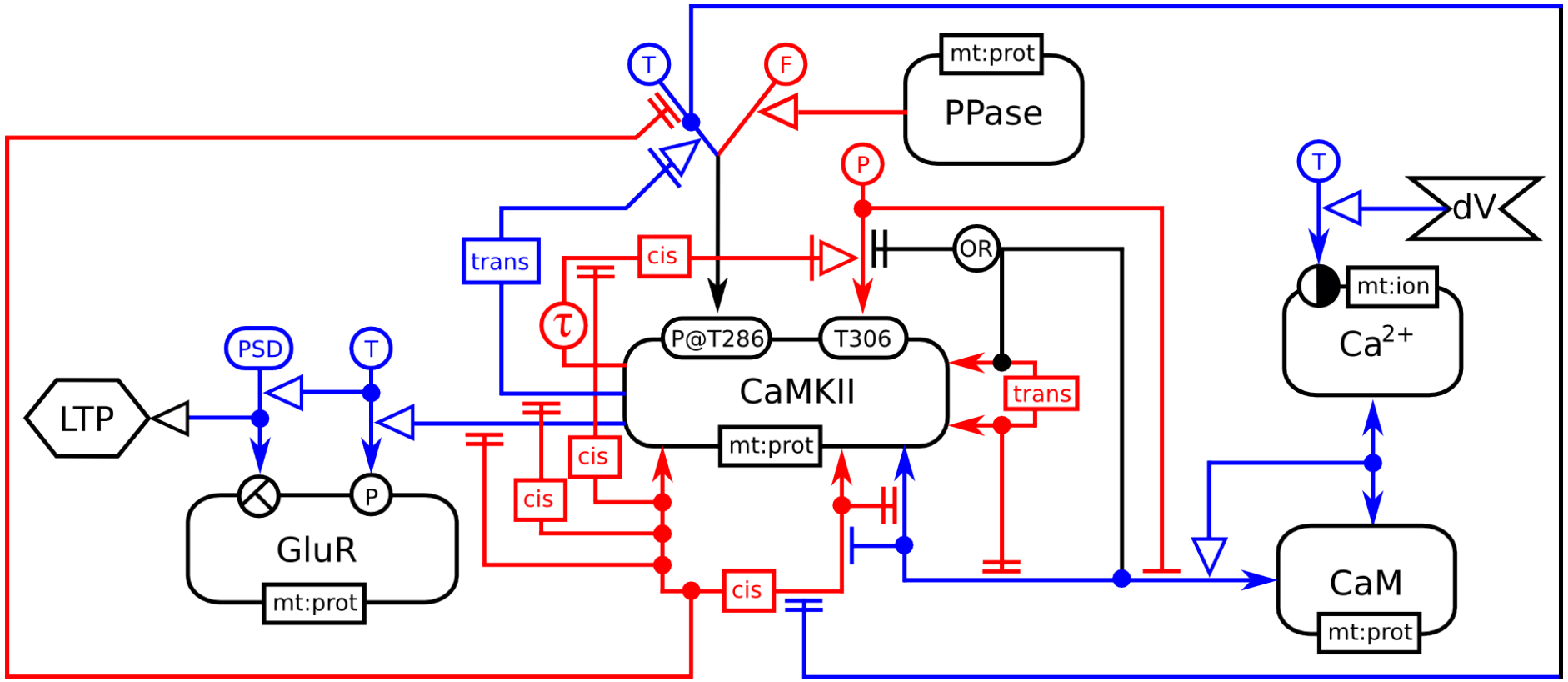


population of multimers



http://sourceforge.net/projects/stochsim/

Regulation of synaptic plasticity by calcium



Biological Pathway Exchange format

http://biopax.org/

Demir *et al* (2010)

BioPAX is a data format and an ontology

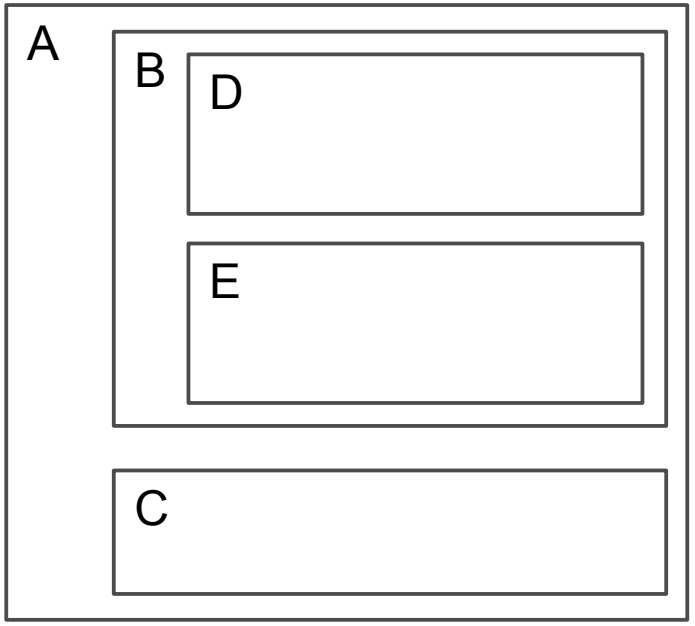
- **RDF/OWL**-based language to represent biological pathways at the molecular and cellular level
- Covers molecular and genetic interactions, metabolic network, signalling pathways, gene regulatory networks
- **Interactions**: Control (Catalysis, Modulation, TemplateReactionRegulation), Conversion (BiochemicalReaction, ComplexAssembly, Degradation, Transport, TransportWithBiochemicalReaction), GeneticInteraction, MolecularInteraction, TemplateReaction
- **Physical entities**: Complex, DNA, DNARegion, Protein, RNA, RNARegion, SmallMolecule
- **Utility classes**: BioSource, ChemicalStructure, ControlledVocabulary, DeltaG, EntityFeature, EntityReference, Evidence, ExperimentalForm, kPrime, PathwayStep, Provenance, Score, SequenceLocation, Stoichiometry, Xref

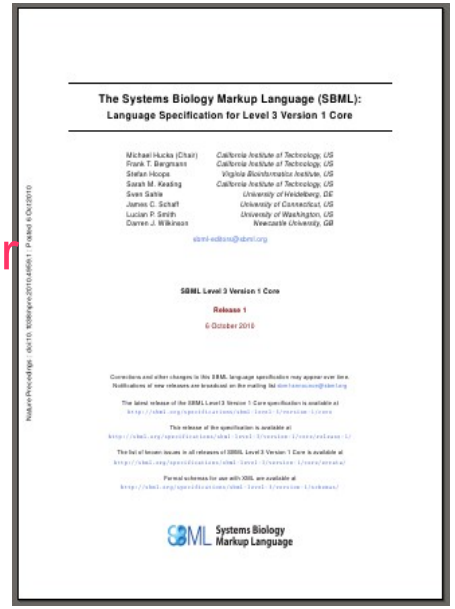
Resource Description Framework - RDF

- Language of the semantic web
- Semantic entirely embedded in the language (does not require external schemas and specifications)
- Series of statements

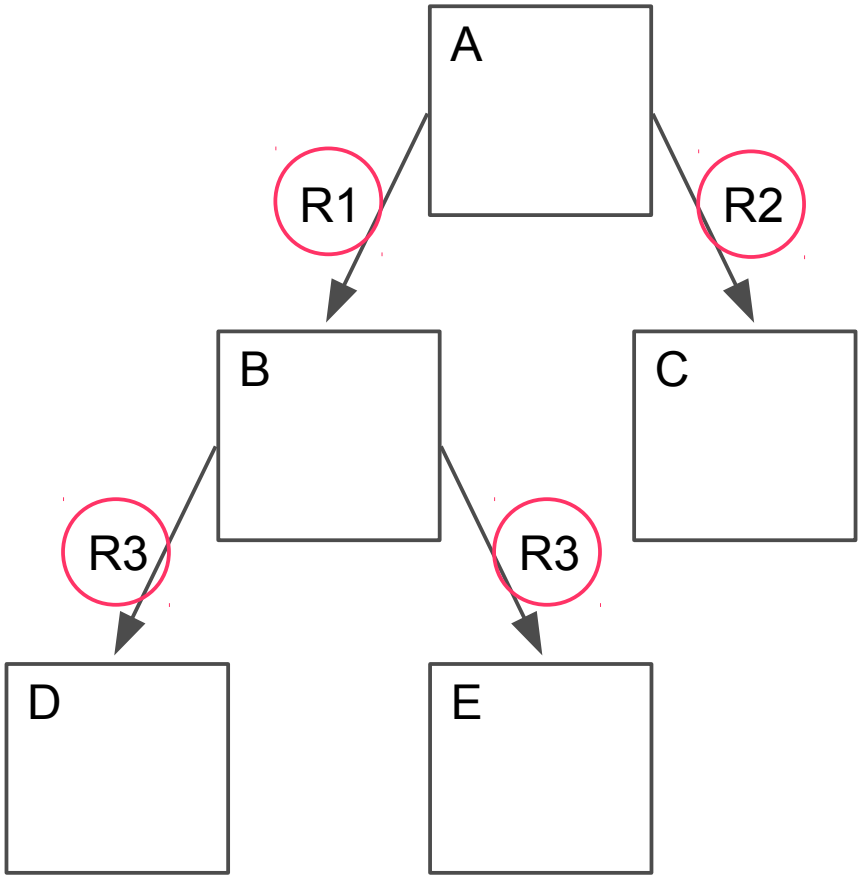
Subject	Predicate	Object
EGFR	*located_in*	plasma membrane
P00533	OBO_REL:0000008	GO:0005886
(UniProt)	(OBO Relation ontology)	(Gene Ontology)

“XML” Vs RDF




The Systems Biology Markup Language (SBML):
Language Specification for Level 3 Version 1 Core

Relationships defined in another document

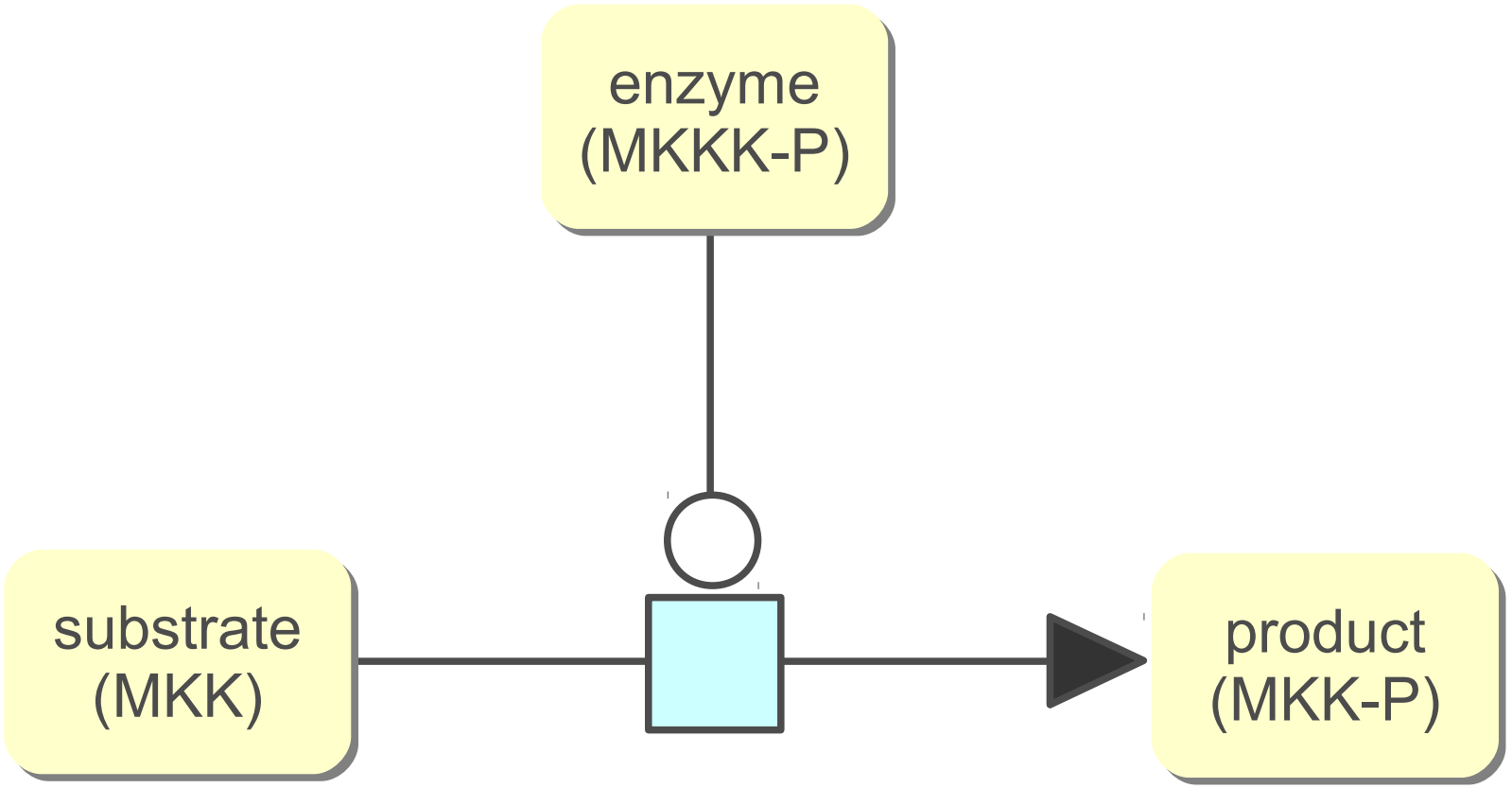


Relationships defined in the pathway

Web Ontology Language - OWL

- BioPAX contains classes, relations and properties, constraints, objects, values.
- Classes are arranged in a specialisation hierarchy. E.g. a “protein” is a “physical entity”
- Classes may have properties of specific types. The types are linked to other classes by relations. E.g. “reaction X” has *participant* “protein P”. An attribute is a property that has a simple type.
- Constraints define allowable values and connexions. E.g. “MOLECULAR_WT” must be a positive real number
- Objects are instances of classes.

A catalysis in BioPAX



A catalysis in BioPAX (L3)

```
<!-- The catalytic reaction -->

<bp:BiochemicalReaction rdf:about="conversion_J2">
  <bp:left rdf:resource="LEFT_0_conversion_J2_MKK" />
  <bp:right rdf:resource="RIGHT_0_conversion_J2_MKK_P" />
</bp:BiochemicalReaction>

<bp:Control rdf:about="control_J2_MKKK_P_0">
 <bp:controlled rdf:resource="conversion_J2" />
 <bp:controlType rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
   ACTIVATION</bp:controlType>
 <bp:controller rdf:resource="PEP0_control_J2MKKK_P" />
</bp:Control>
```

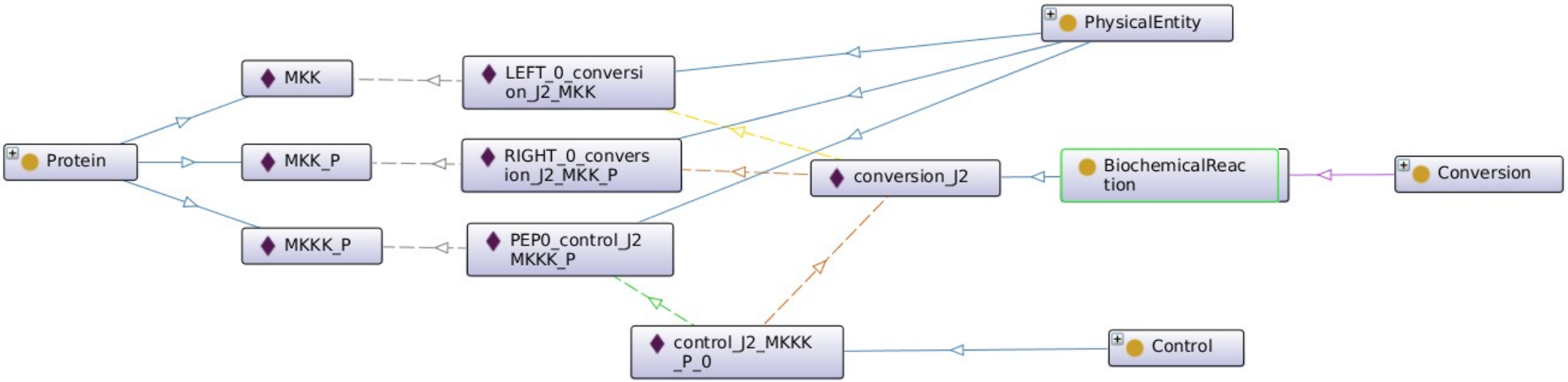
A catalysis in BioPAX (L3)

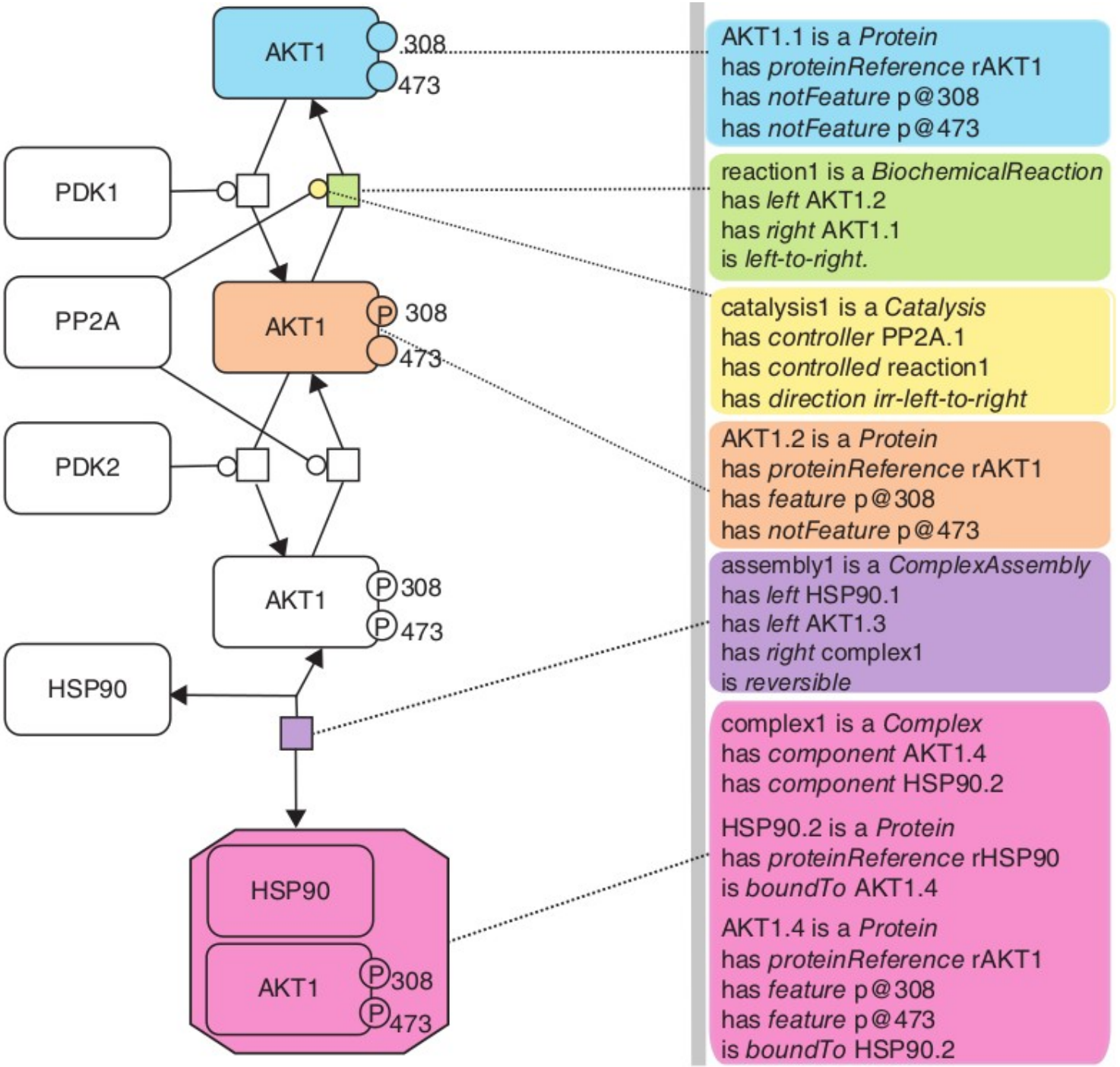
```
<!-- The participants -->

<bp:PhysicalEntity rdf:about="RIGHT_0_conversion_J2_MKK_P">
  <bp:displayName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
  MKK_P</bp:displayName>
  <bp:memberPhysicalEntity rdf:resource="MKK_P" />
</bp:PhysicalEntity>

<bp:PhysicalEntity rdf:about="LEFT_0_conversion_J2_MKK">
  <bp:displayName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
  MKK</bp:displayName>
  <bp:memberPhysicalEntity rdf:resource="MKK" />
</bp:PhysicalEntity>

<bp:PhysicalEntity rdf:about="PEP0_control_J2MKKK_P">
  <bp:displayName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
  MKKK_P</bp:displayName>
  <bp:memberPhysicalEntity rdf:resource="MKKK_P" />
</bp:PhysicalEntity>
```

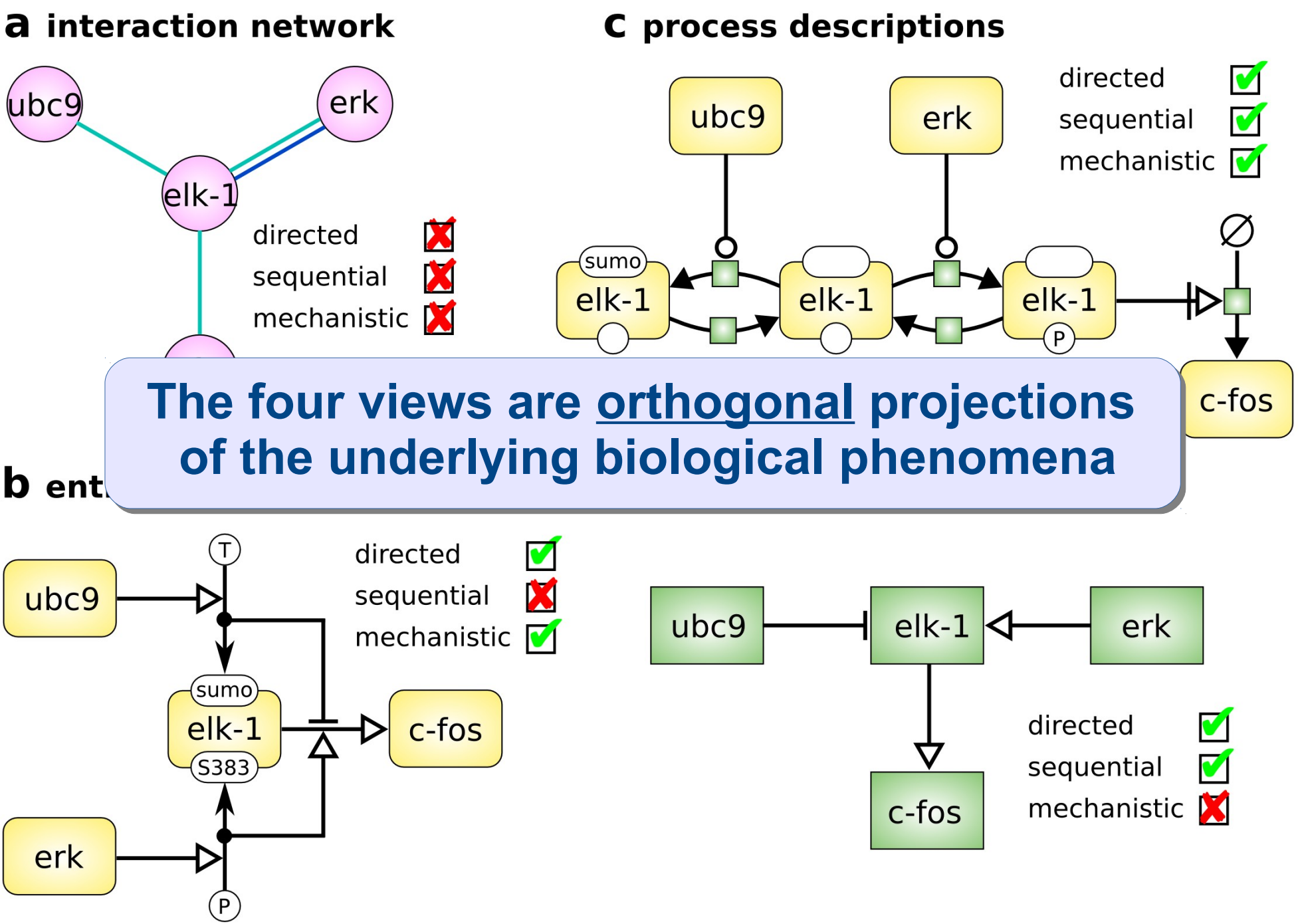
A catalysis in BioPAX (L3)

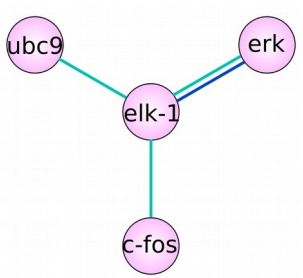
```
<!-- Proteins -->

<bp:Protein rdf:about="MKKK_P">
 <bp:displayName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
   Mos-P</bp:displayName>
</bp:Protein>

<bp:Protein rdf:about="MKK">
 <bp:displayName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
   Mek1</bp:displayName>
</bp:Protein>

<bp:Protein rdf:about="MKK_P">
 <bp:displayName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
   Mek1-P</bp:displayName>
</bp:Protein>
```

RDF graph (Protégé, http://protege.stanford.edu/)



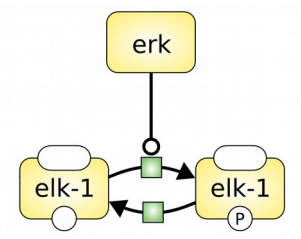

AKT1
308
473
PDK1
PP2A
AKT1
P
308
473
PDK2
AKT1
P
308
P
473
HSP90
HSP90
AKT1
P
308
P
473
AKT1.1 is a Protein
has proteinReference rAKT1
has notFeature p@308
has notFeature p@473
reaction1 is a BiochemicalReaction
has left AKT1.2
has right AKT1.1
is left-to-right.
catalysis1 is a Catalysis
has controller PP2A.1
has controlled reaction1
has direction irr-left-to-right
AKT1.2 is a Protein
has proteinReference rAKT1
has feature p@308
has notFeature p@473
assembly1 is a ComplexAssembly
has left HSP90.1
has left AKT1.3
has right complex1
is reversible
complex1 is a Complex
has component AKT1.4
has component HSP90.2
HSP90.2 is a Protein
has proteinReference rHSP90
is boundTo AKT1.4
AKT1.4 is a Protein
has proteinReference rAKT1
has feature p@308
has feature p@473
is boundTo HSP90.2


a interaction network
ubc9
erk
elk-1
directed
sequential
mechanistic
c process descriptions
ubc9
erk
directed
sequential
mechanistic
sumo
elk-1
elk-1
elk-1
P
c-fos
The four views are orthogonal projections of the underlying biological phenomena
b ent
T
ubc9
directed
sequential
mechanistic
sumo
elk-1
S383
c-fos
erk
P
ubc9
elk-1
erk
c-fos
directed
sequential
mechanistic



Interaction networks

PSI-MI

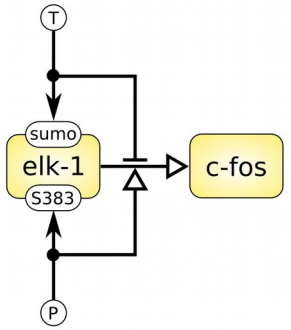


Process Descriptions

SBML Core

SBGN Process Description

Metabolism

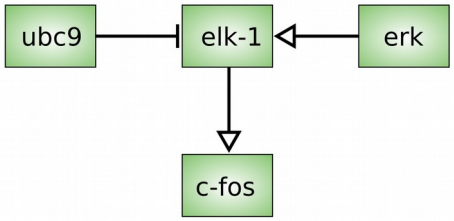


Entity relationships

SBML Multi

SBGN Entity Relationships

Signalling



Activity Flows

SBML Qual

SBGN Activity Flows

Signalling
Gene networks