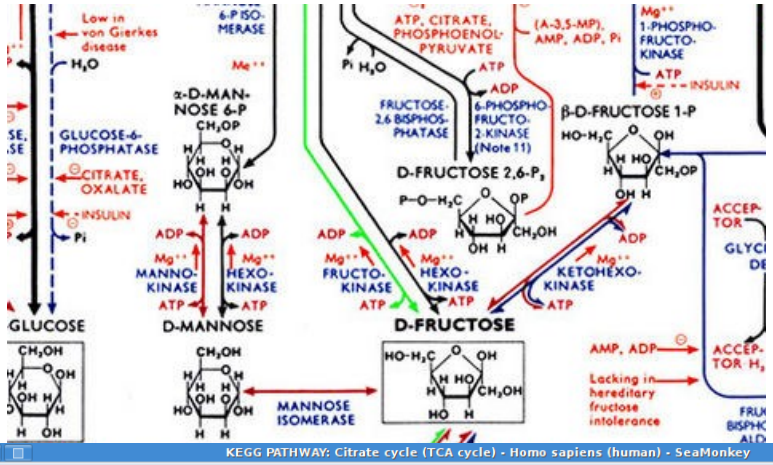
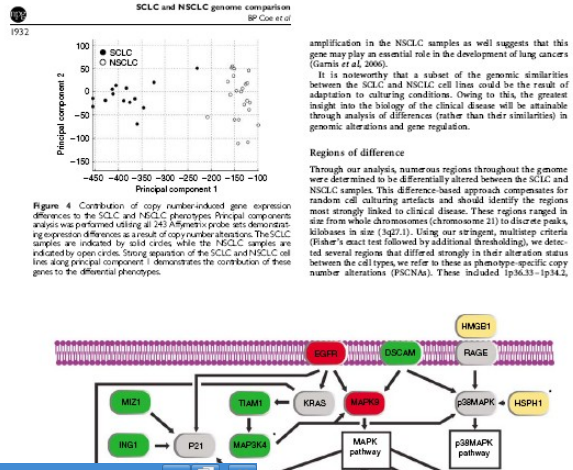
The Systems Biology Graphical Notation

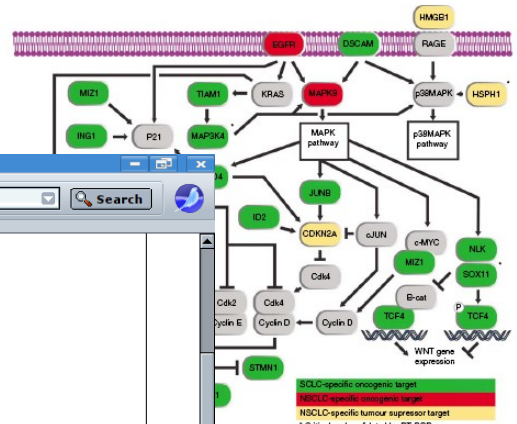
Nicolas Le Novère, EMBL-EBI

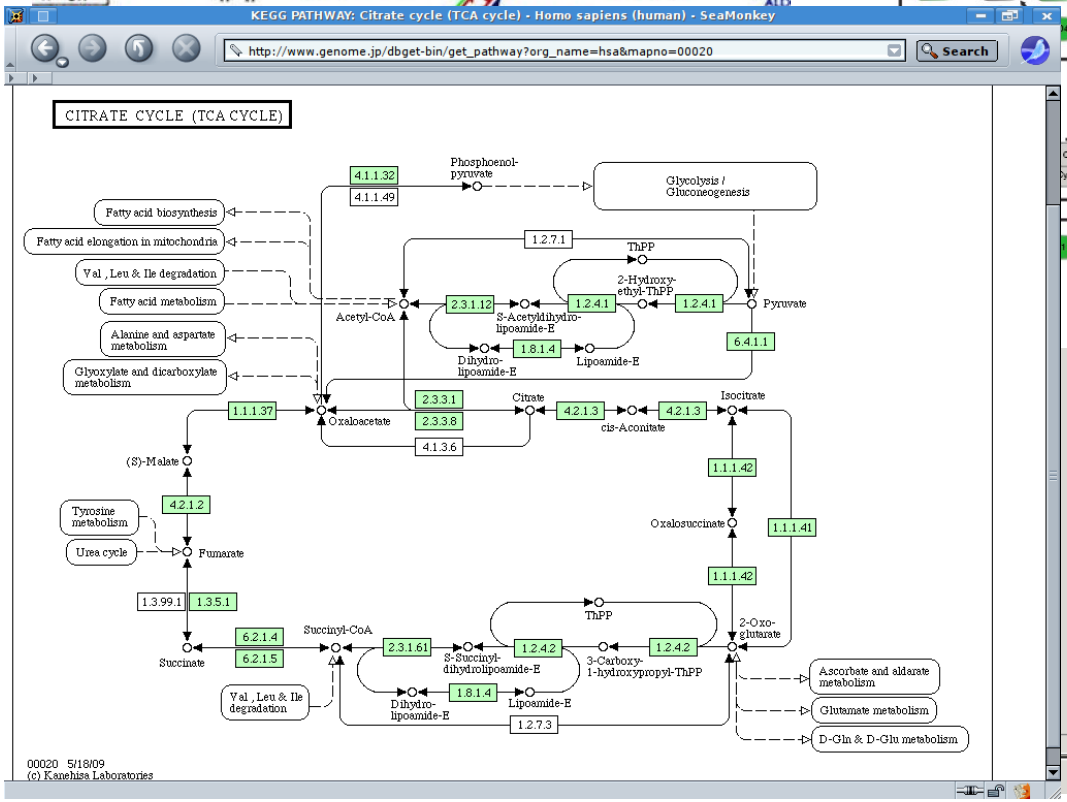
(on the behalf of SBGN editors, authors and contributors)

Graphs are everywhere in biology

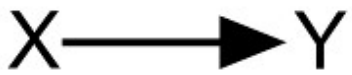




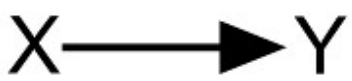




Ambiguity of usual representations



Ambiguity of usual representations



is transformed into

translocates (X "=" Y)

is degraded into

associates into

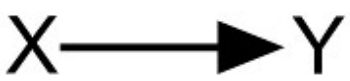
dissociates into

stimulates the activity of

stimulates the expression of

catalyses the formation of

Ambiguity of usual representations



X inhibits Y

is transformed into

translocates (X "=" Y)

is degraded into

associates into

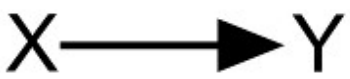
dissociates into

stimulates the activity of

stimulates the expression of

catalyses the formation of

Ambiguity of usual representations



is transformed into

translocates (X "=" Y)

is degraded into

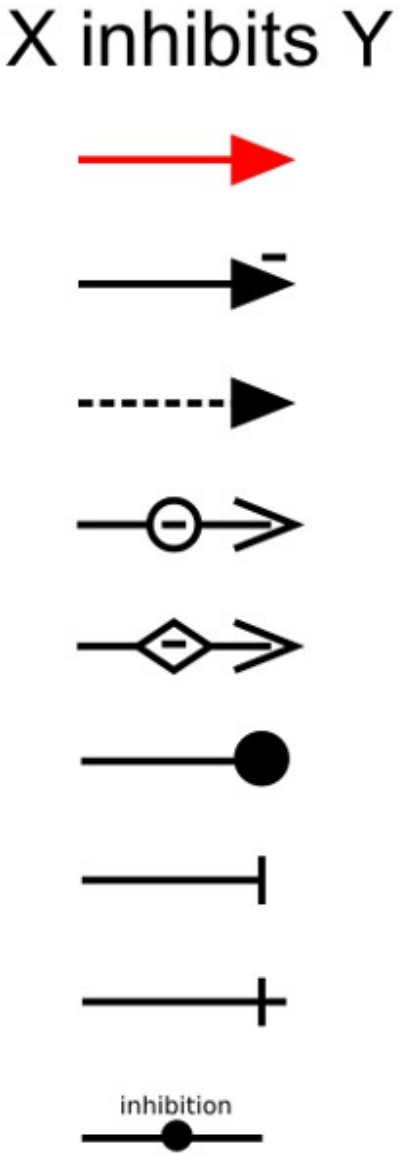
associates into

dissociates into

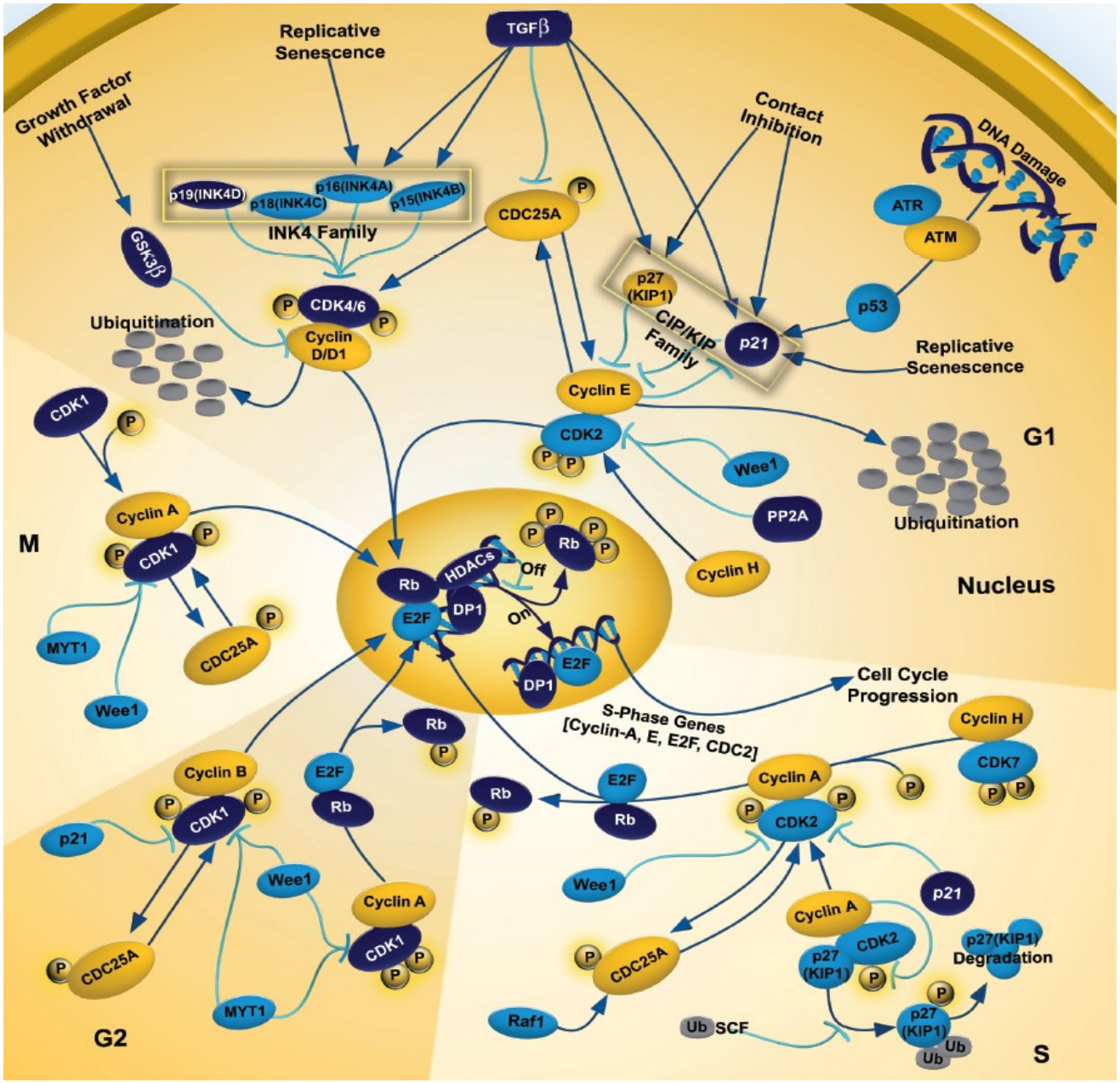
stimulates the activity of

stimulates the expression of

catalyses the formation of



Can-this be understood by biologists?



Can-this be understood by biologists?

Stimulates? but … what exactly?

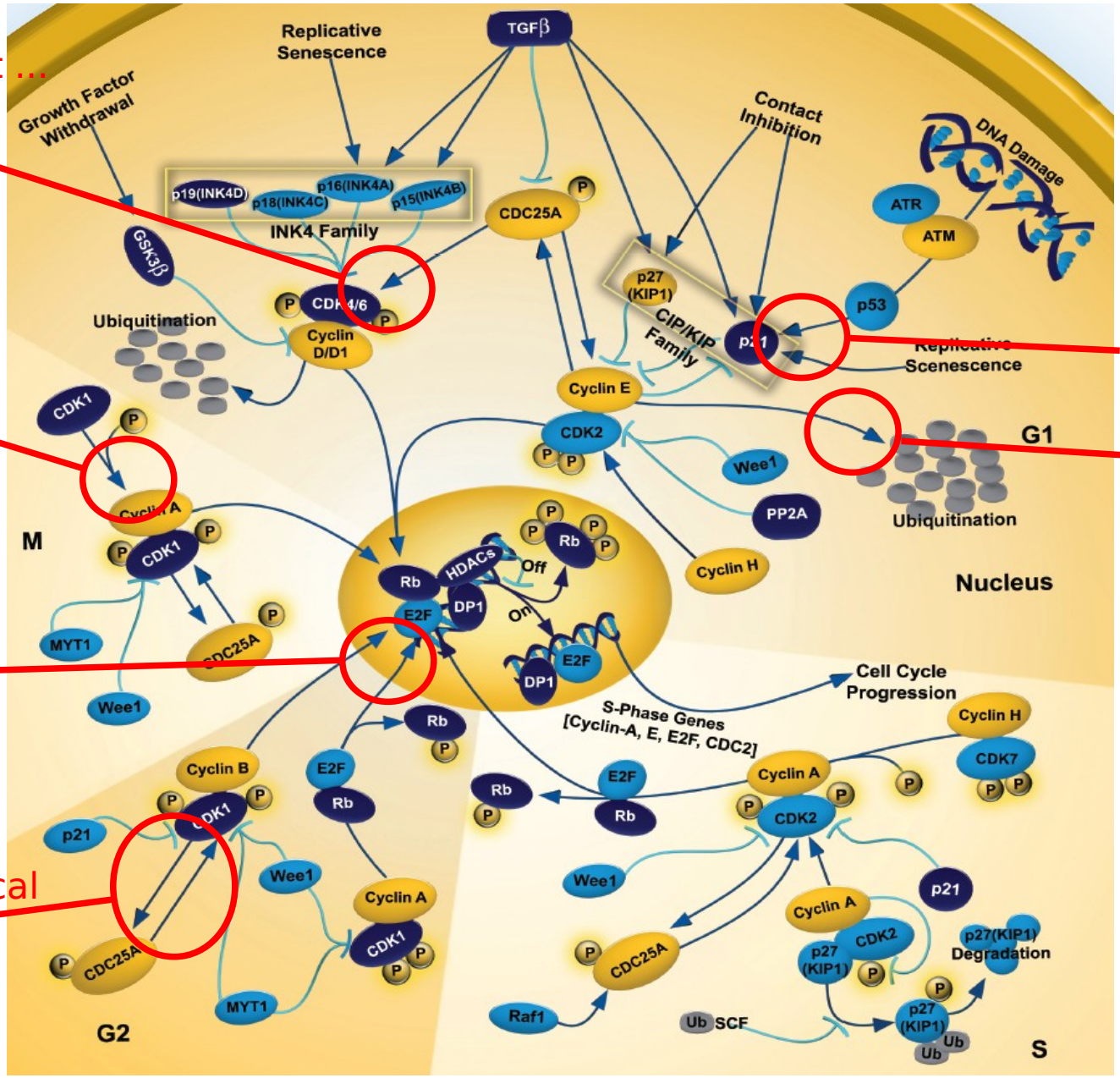
Associates into?

Translocates?

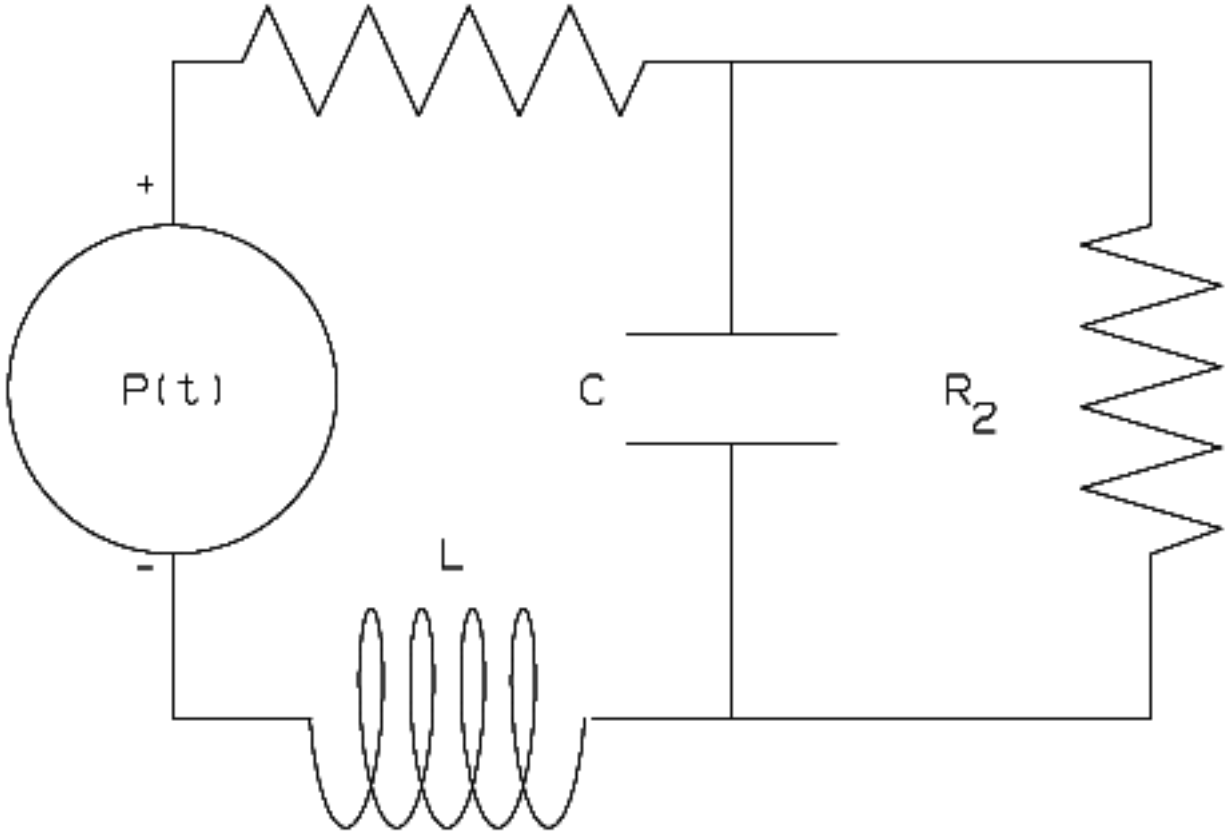
No idea. Reciprocal stimulation?

Stimulates gene transcription?

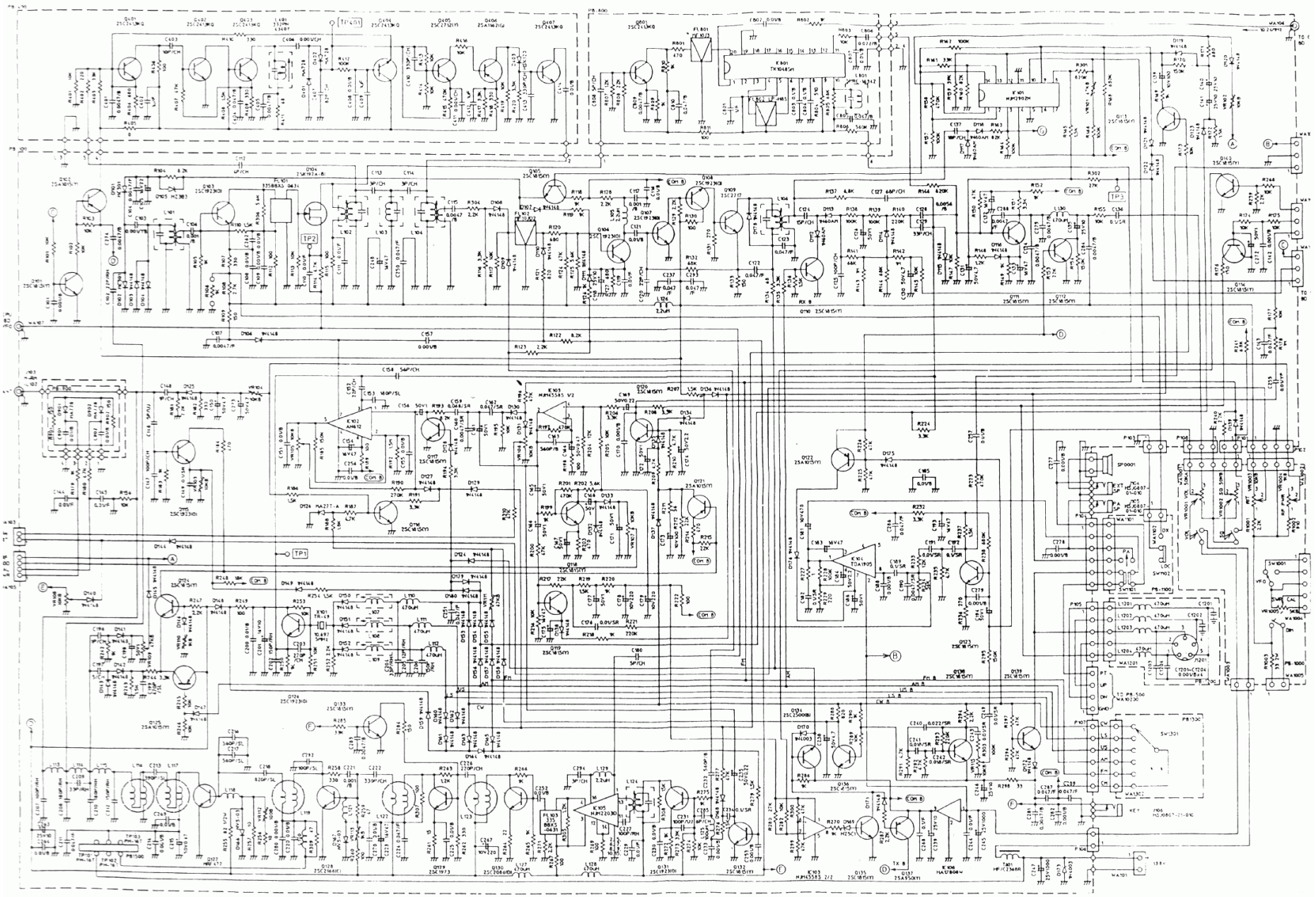
Is degraded?



Every engineer understands that



Or that



Why can-we understand electric diagrams?

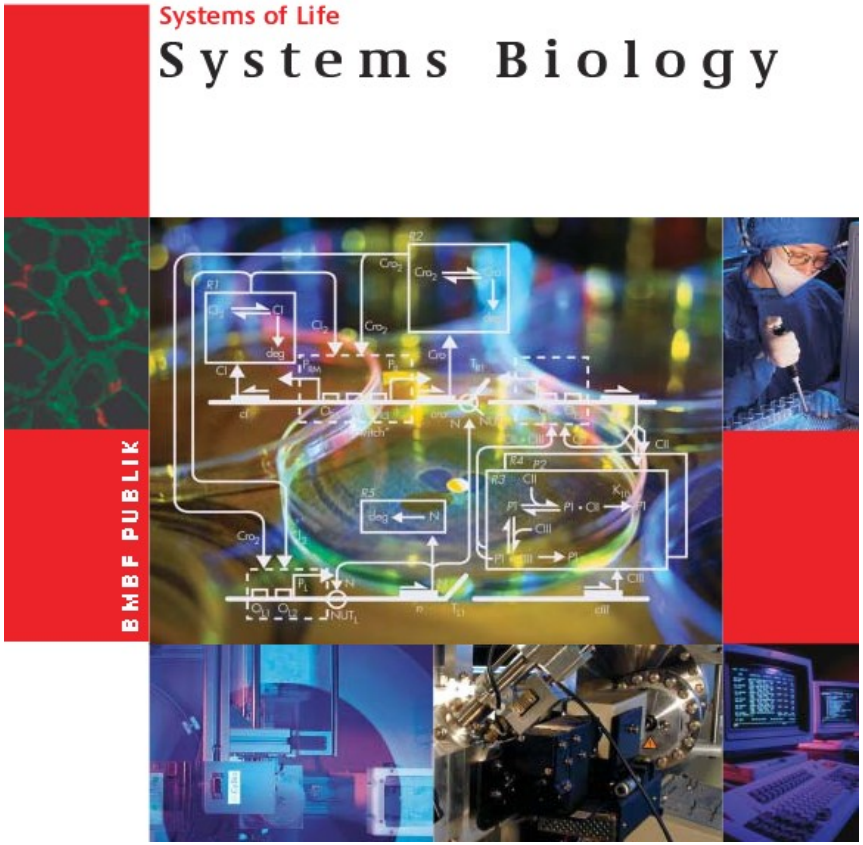
- Standard symbols
  - Simple shapes, easily recognisable
  - Limited number of basic symbols (<70)
  - Similarity of shapes reflects similarity of functions
- Unambiguous interpretation of the circuits
- Endorsed by the community for practical reasons
  - End-users: manufacturers
  - Tool developers
  - Publishing industry
  - Teaching communities

What did-those diagram bring?

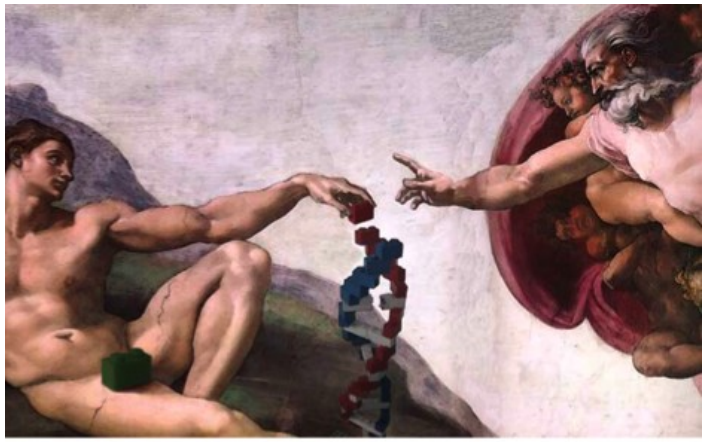
What do-we expect in modern (future) life science

Federal Ministry
of Education
and Research

Basic science



Technology



EXTREME GENETIC ENGINEERING

An Introduction to Synthetic Biology

etc group

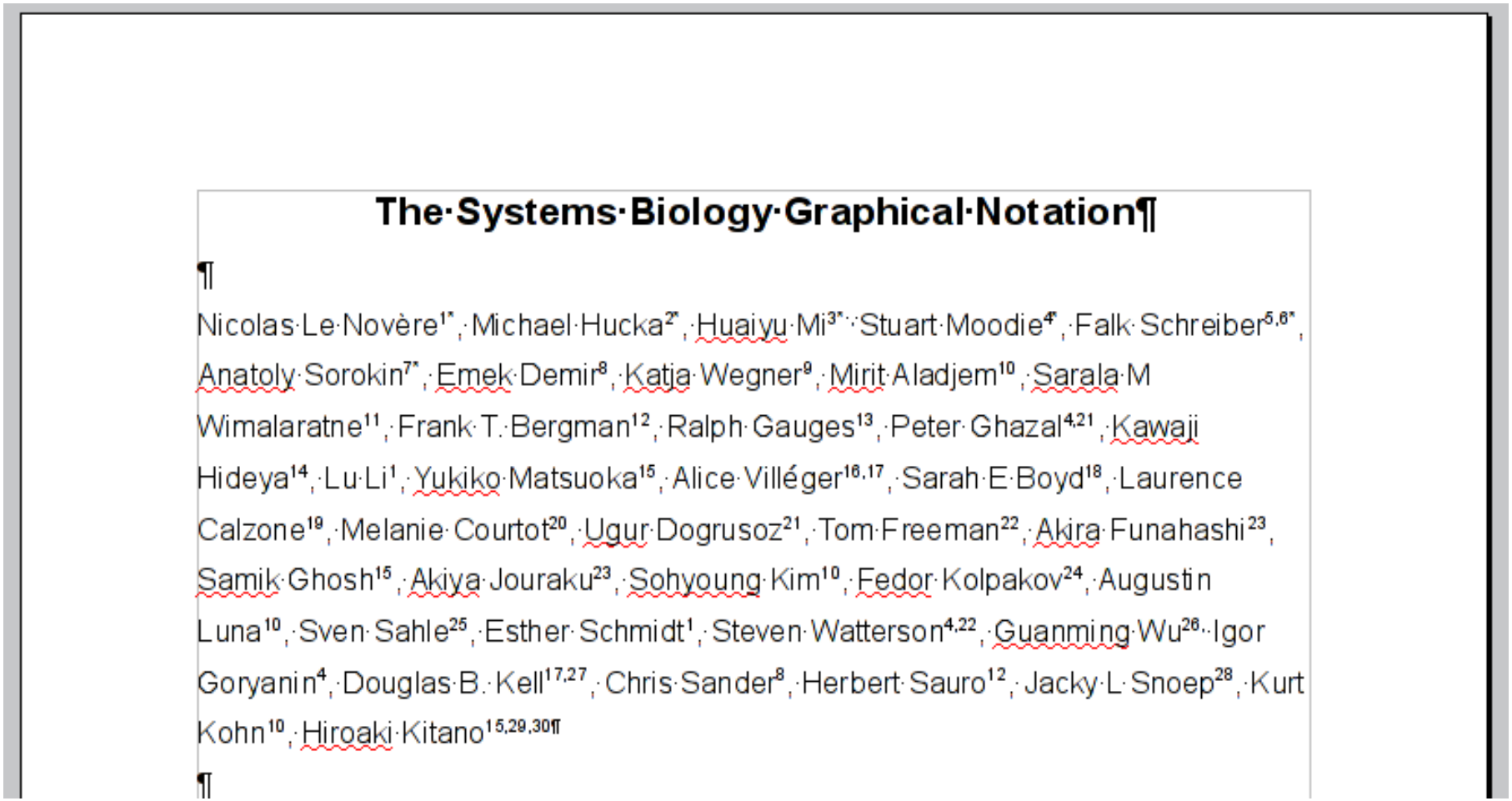
January 2007

Enters
The Systems Biology Markup Language

http://www.sbgn.org/

What is SBGN?

- A way to unambiguously describe biochemical and cellular events in graphs
- Limited amount of symbols ☞ Smooth learning curve
- Can graphically represent quantitative models, biochemical pathways, at different levels of granularity
- Developed over three years by a growing community



The·Systems·Biology·Graphical·Notation¶

¶

Nicolas·Le·Novère[1*],·Michael·Hucka[2*],·Huaiyu·Mi[3*],·Stuart·Moodie[4*],·Falk·Schreiber[5,6*],·Anatoly·Sorokin[7*],·Emek·Demir[8],·Katja·Wegner[9],·Mirit·Aladjem[10],·Sarala·M Wimalaratne[11],·Frank·T.·Bergman[12],·Ralph·Gauges[13],·Peter·Ghazal[4,21],·Kawaji Hideya[14],·Lu·Li[1],·Yukiko·Matsuoka[15],·Alice·Villéger[16,17],·Sarah·E·Boyd[18],·Laurence Calzone[19],·Melanie·Courtot[20],·Ugur·Dogrusoz[21],·Tom·Freeman[22],·Akira·Funahashi[23],·Samik·Ghosh[15],·Akiya·Jouraku[23],·Sohyoung·Kim[10],·Fedor·Kolpakov[24],·Augustin Luna[10],·Sven·Sahle[25],·Esther·Schmidt[1],·Steven·Watterson[4,22],·Guanming·Wu[26],·Igor Goryanin[4],·Douglas·B.·Kell[17,27],·Chris·Sander[8],·Herbert·Sauro[12],·Jacky·L·Snoep[28],·Kurt Kohn[10],·Hiroaki·Kitano[15,29,30]¶

¶

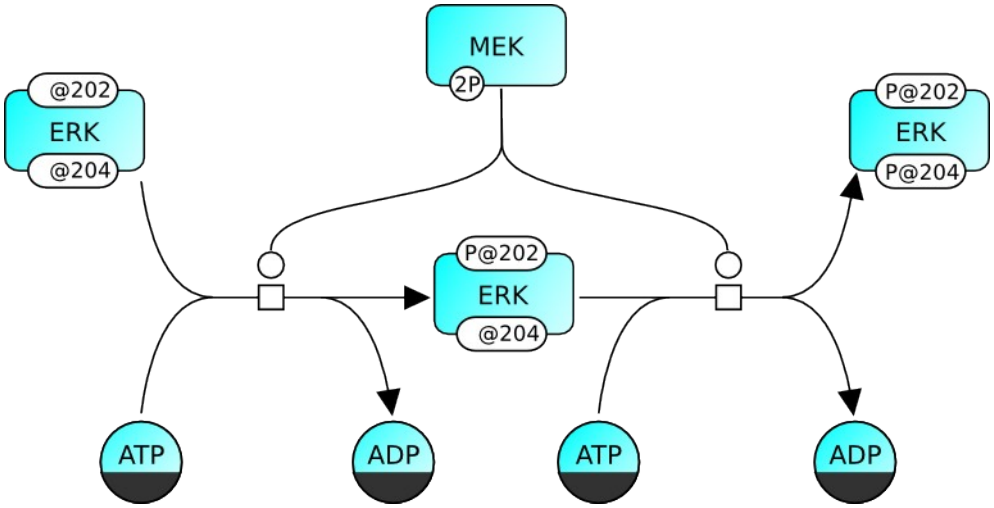
39 authors, 30 affiliations

What is SBGN?

- A way to unambiguously describe biochemical and cellular events in graphs
- Limited amount of symbols ☞ Smooth learning curve
- Can graphically represent quantitative models, biochemical pathways, at different levels of granularity
- Developed over three years by a growing community
- Three languages
  - Process Diagrams ☞ one state = one glyph, biochemical level
  - Entity Relationships ☞ one entity = one glyph, biochemical level
  - Activity Flow ☞ conceptual level

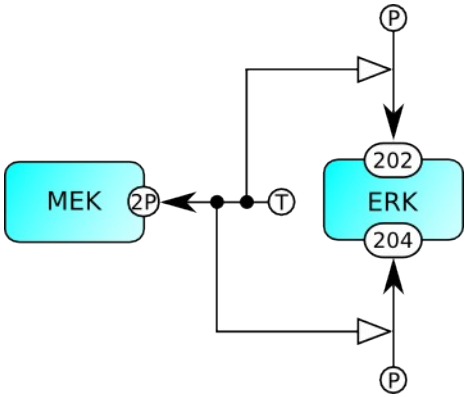
Graph trinity: three languages in one

Process diagrams



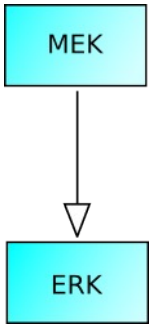
- Unambiguous
- Mechanistic
- Sequential
- Subject to combinatorial explosion

Entity Relationships diagrams



- Unambiguous
- Mechanistic
- Non-sequential

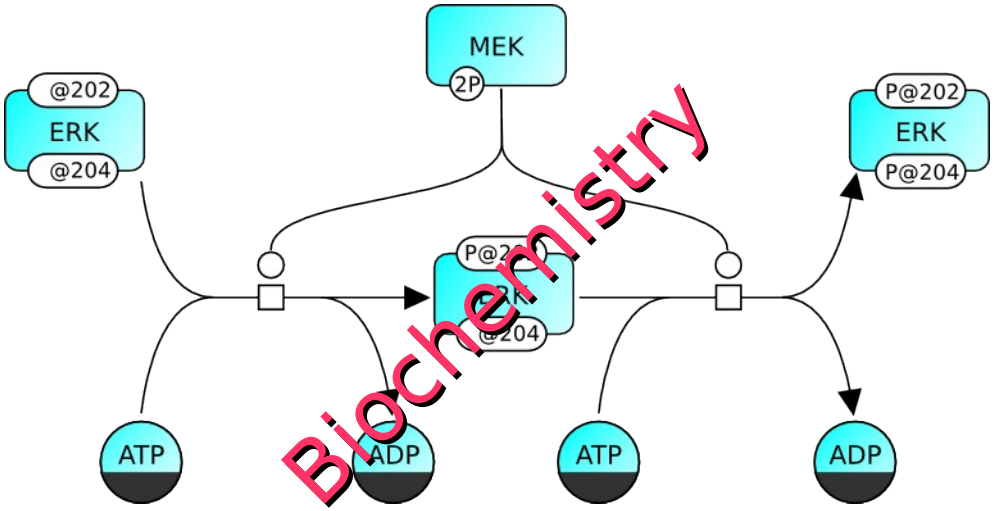
Activity Flow diagrams



- Ambiguous
- Conceptual
- Sequential

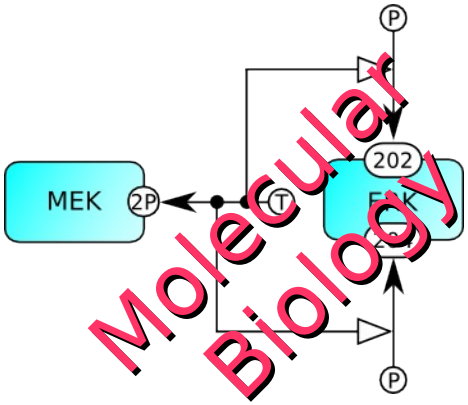
Graph trinity: three languages in one

Process diagrams



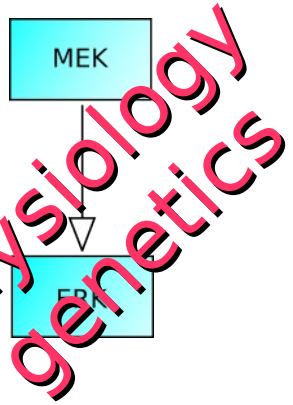
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

Entity Relationships diagrams



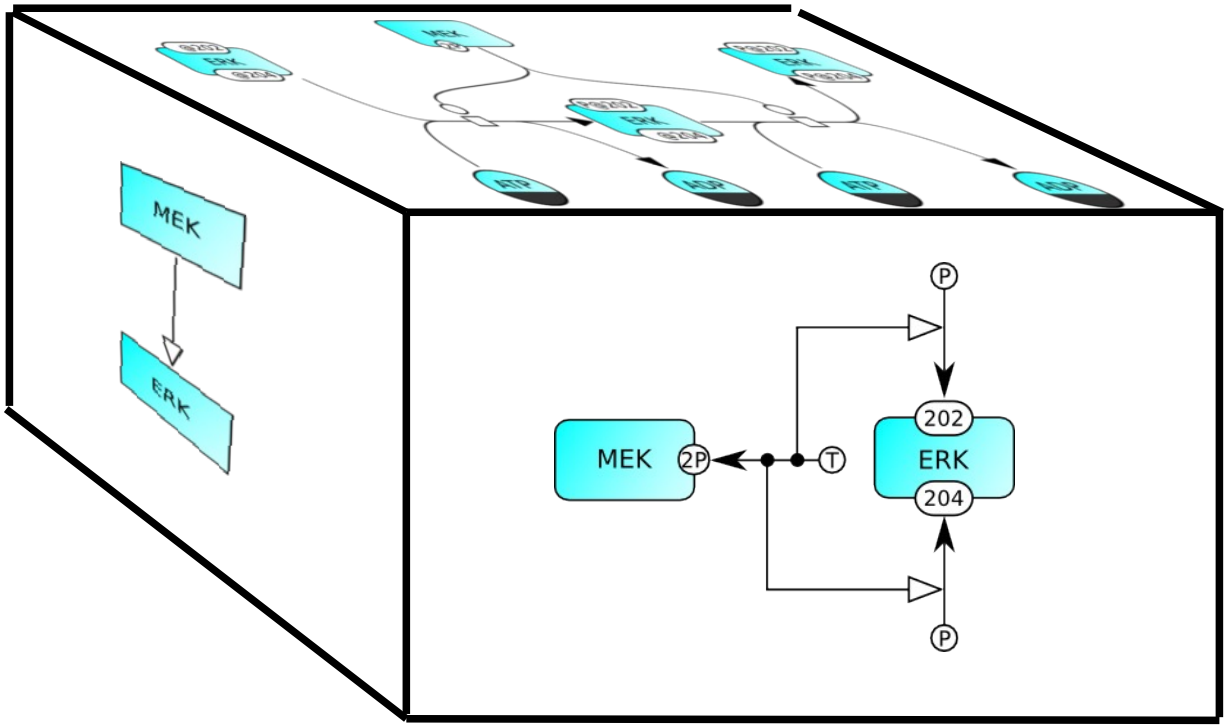
- Unambiguous
- Mechanistic
- Non-sequential
- Independence of relationships

Activity Flow diagrams

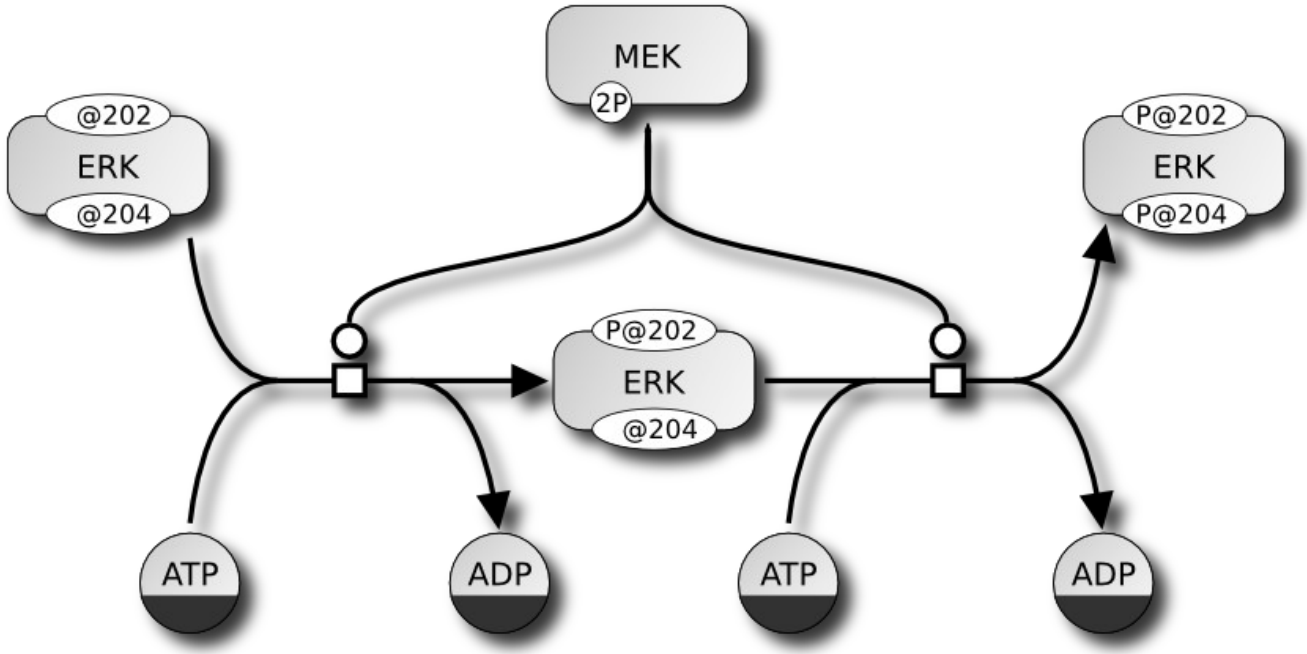


- Ambiguous
- Conceptual
- Sequential

Three orthogonal projections of biology

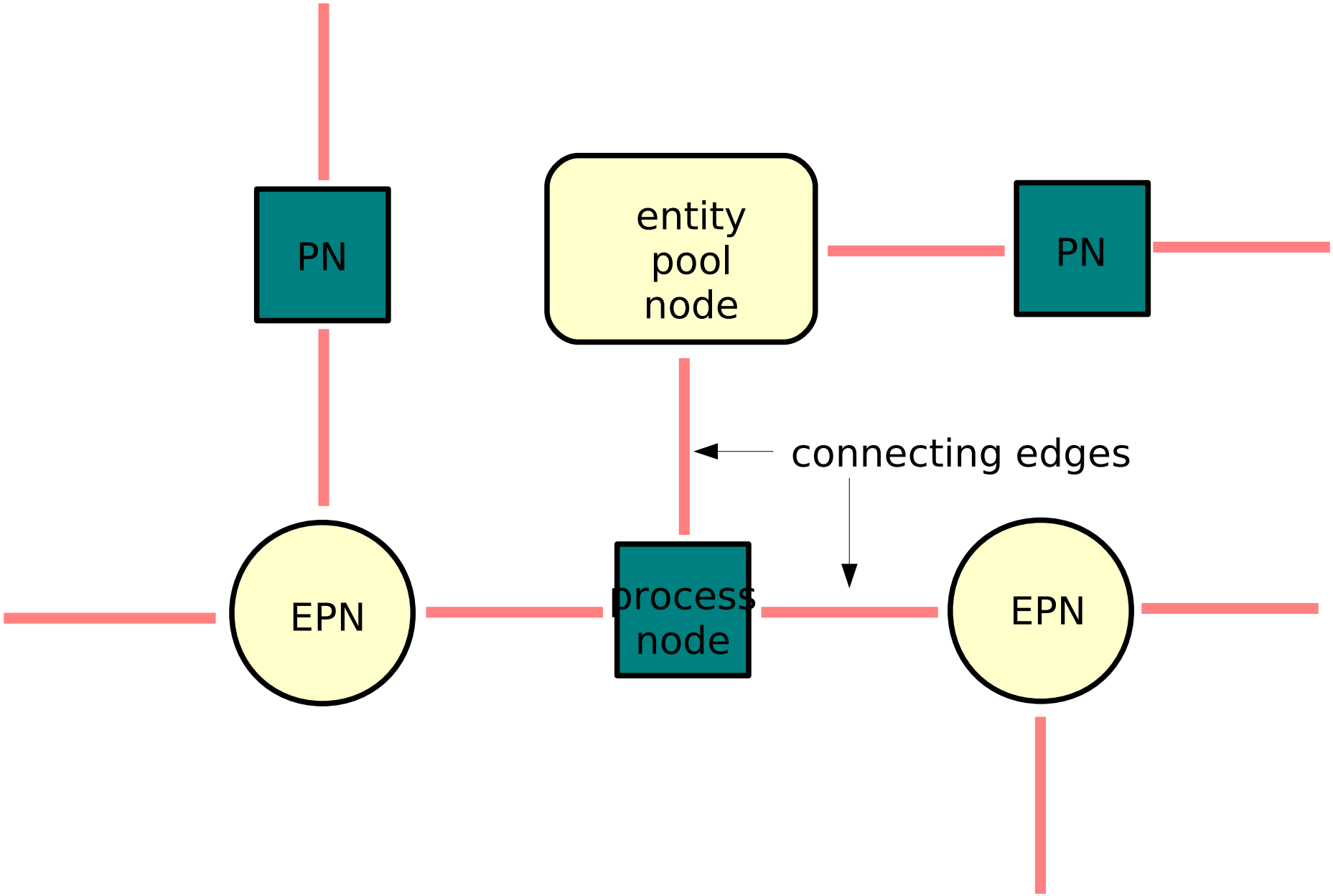


SBGN Process Diagram Level 1

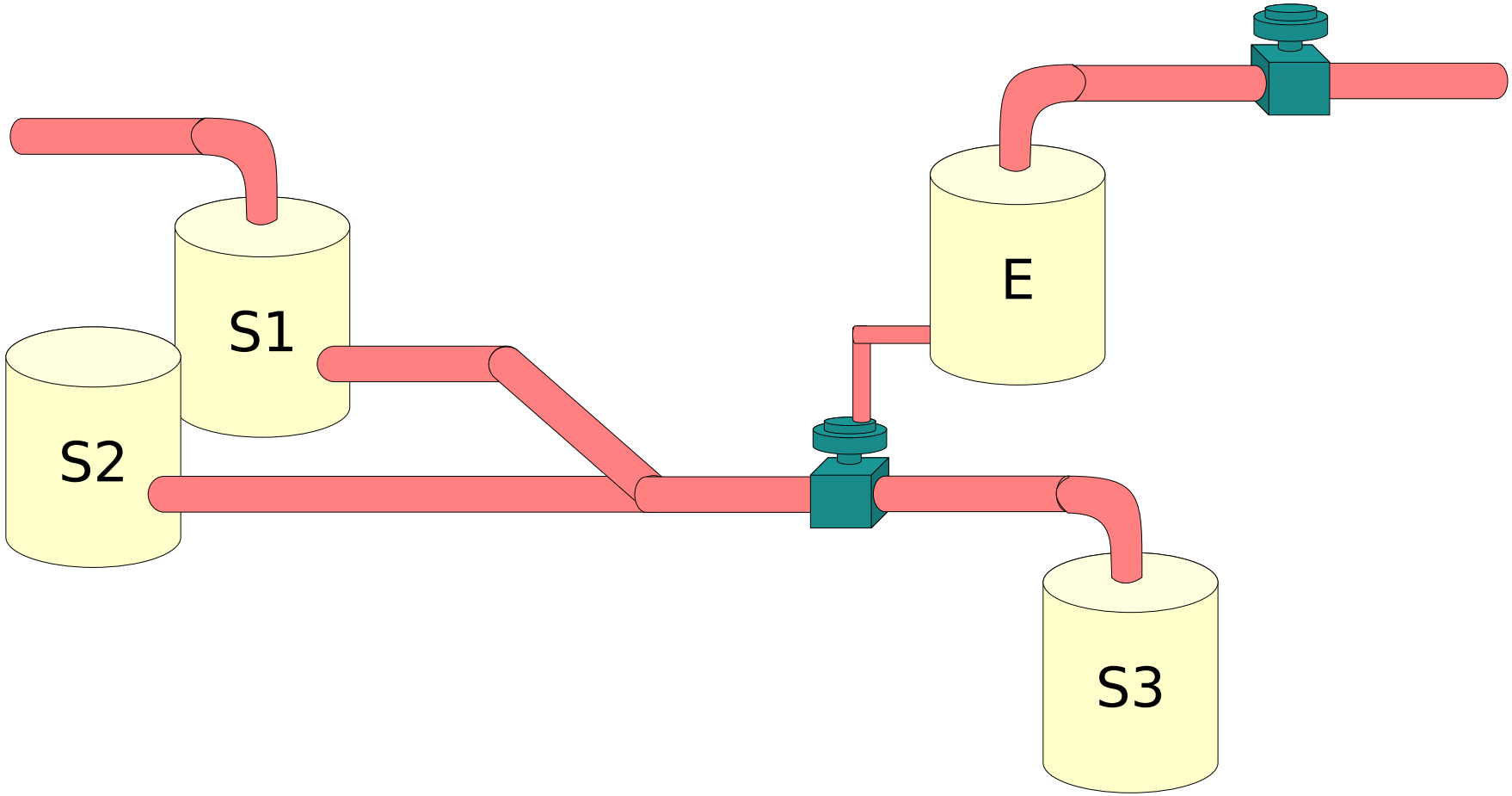


Le Novère, Moodie, Sorokin, Hucka, Schreiber, Demir, Mi, Matsuoka, Wegner, Kitano
Systems Biology Graphical Notation: Process Diagram Level 1 (2008)
Available from *Nature Precedings* <http://hdl.handle.net/10101/npre.2008.2320.1>

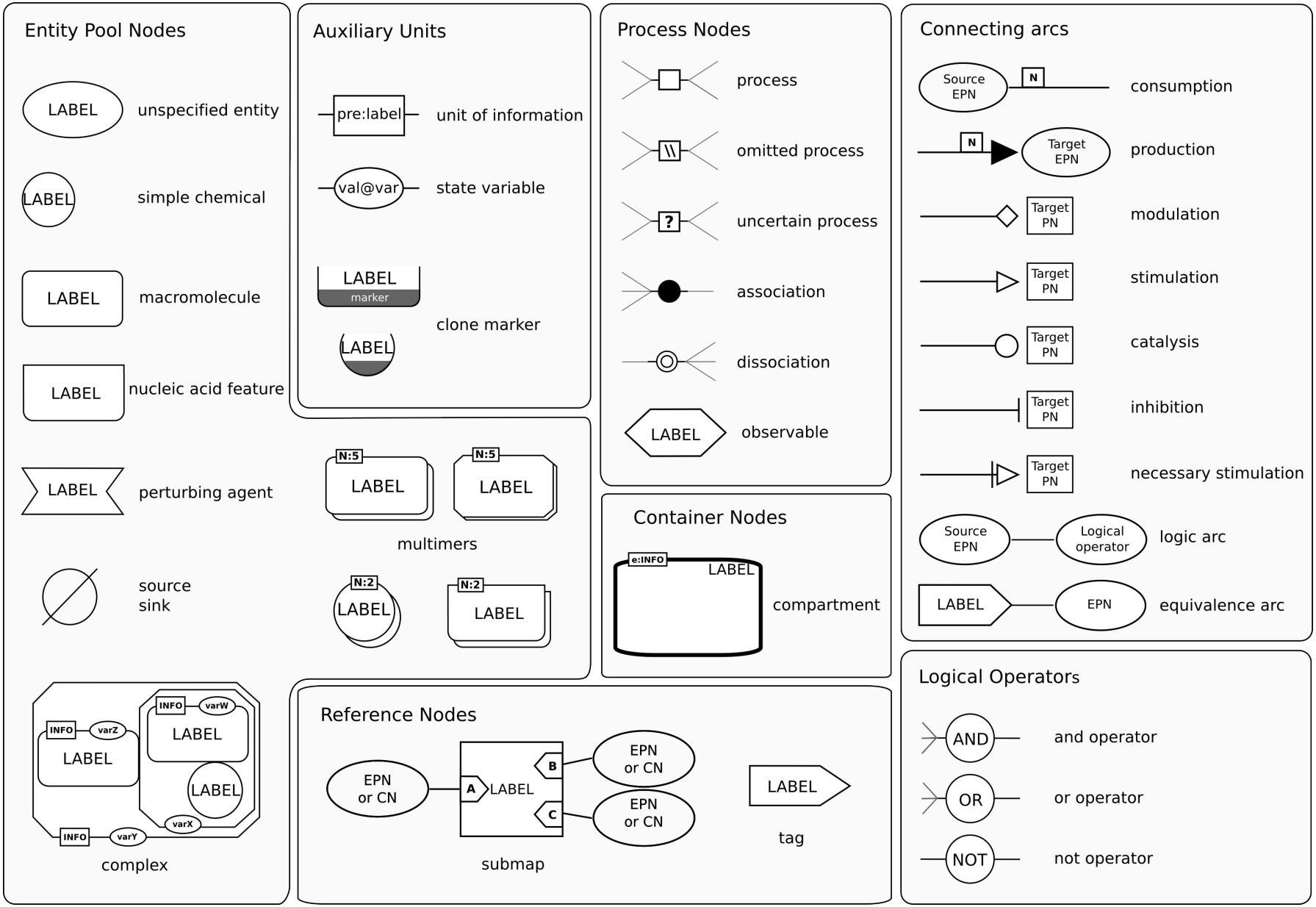
Process Diagrams are bipartite graphs



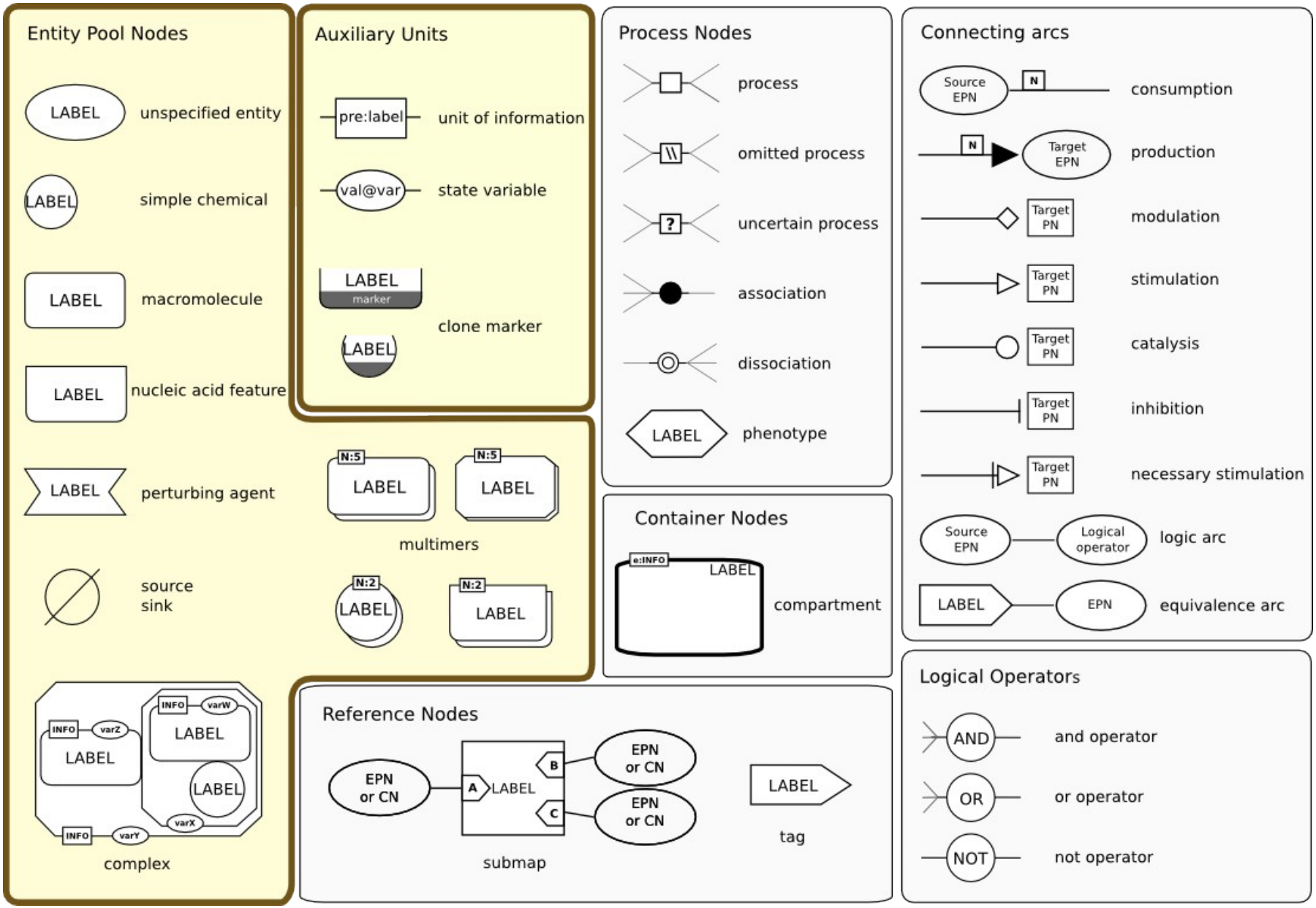
Process Diagrams can be viewed as pipelines



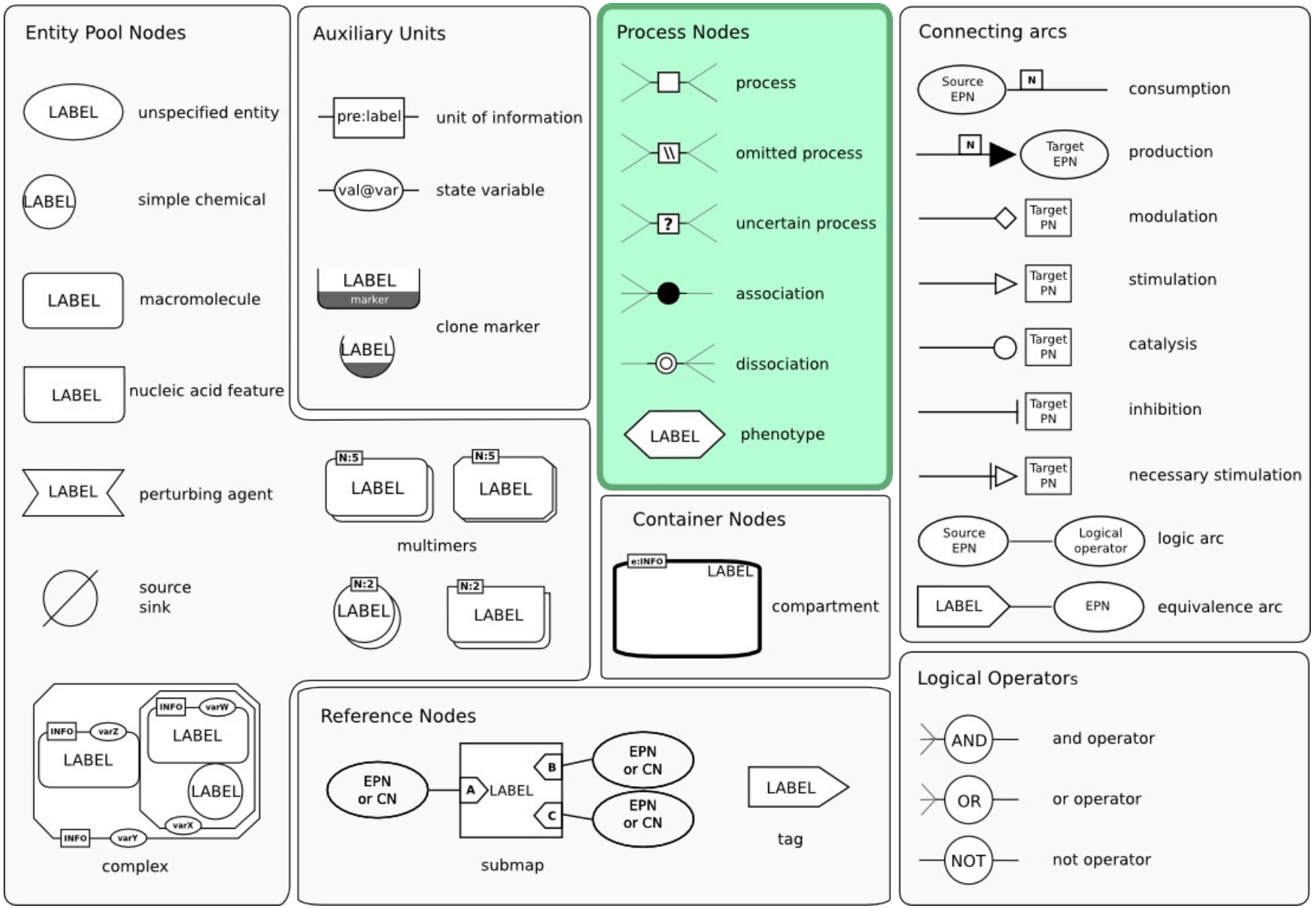
SBGN Process Diagram L1 reference card



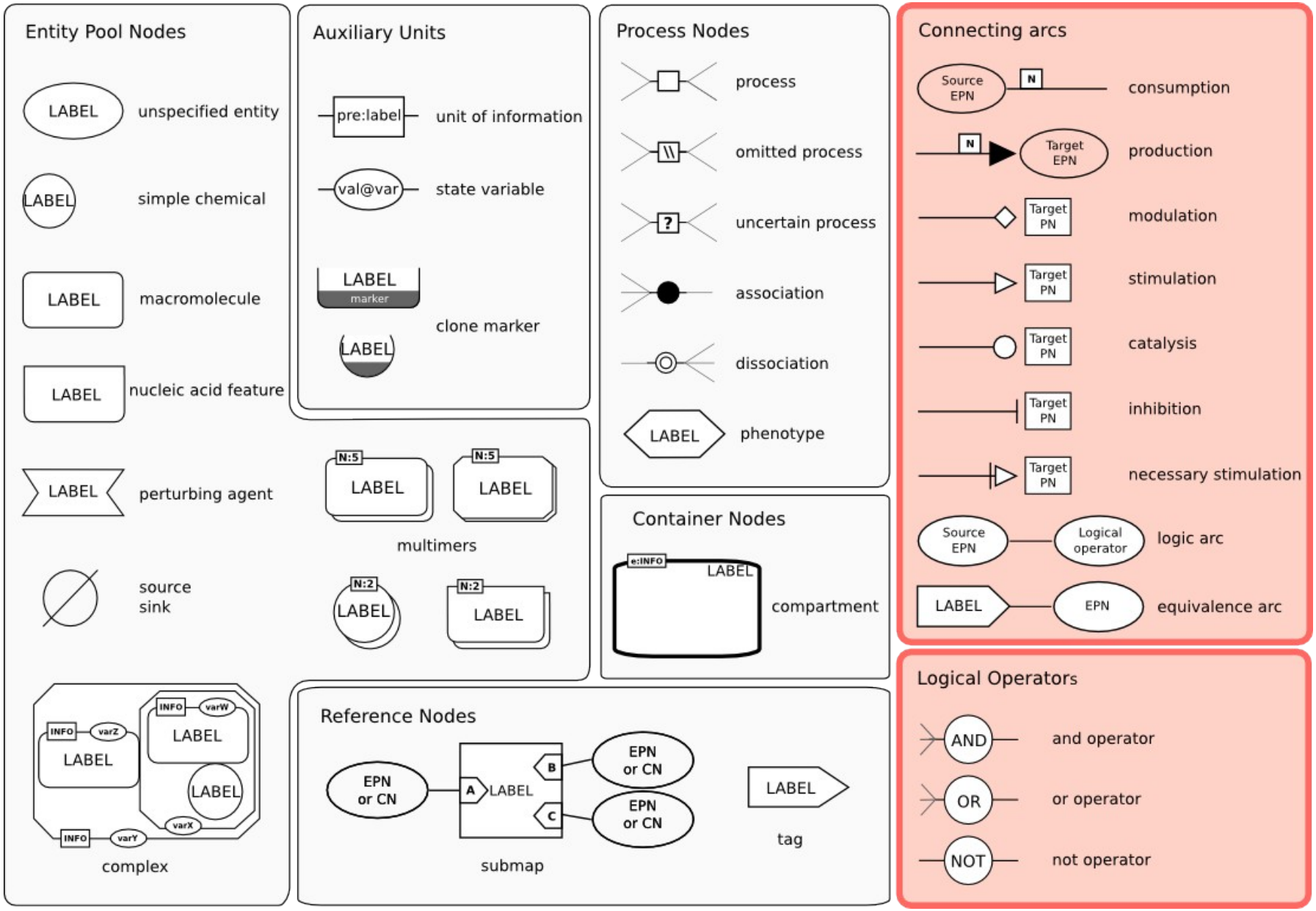
Entity Pool Nodes



Process Nodes



Edges

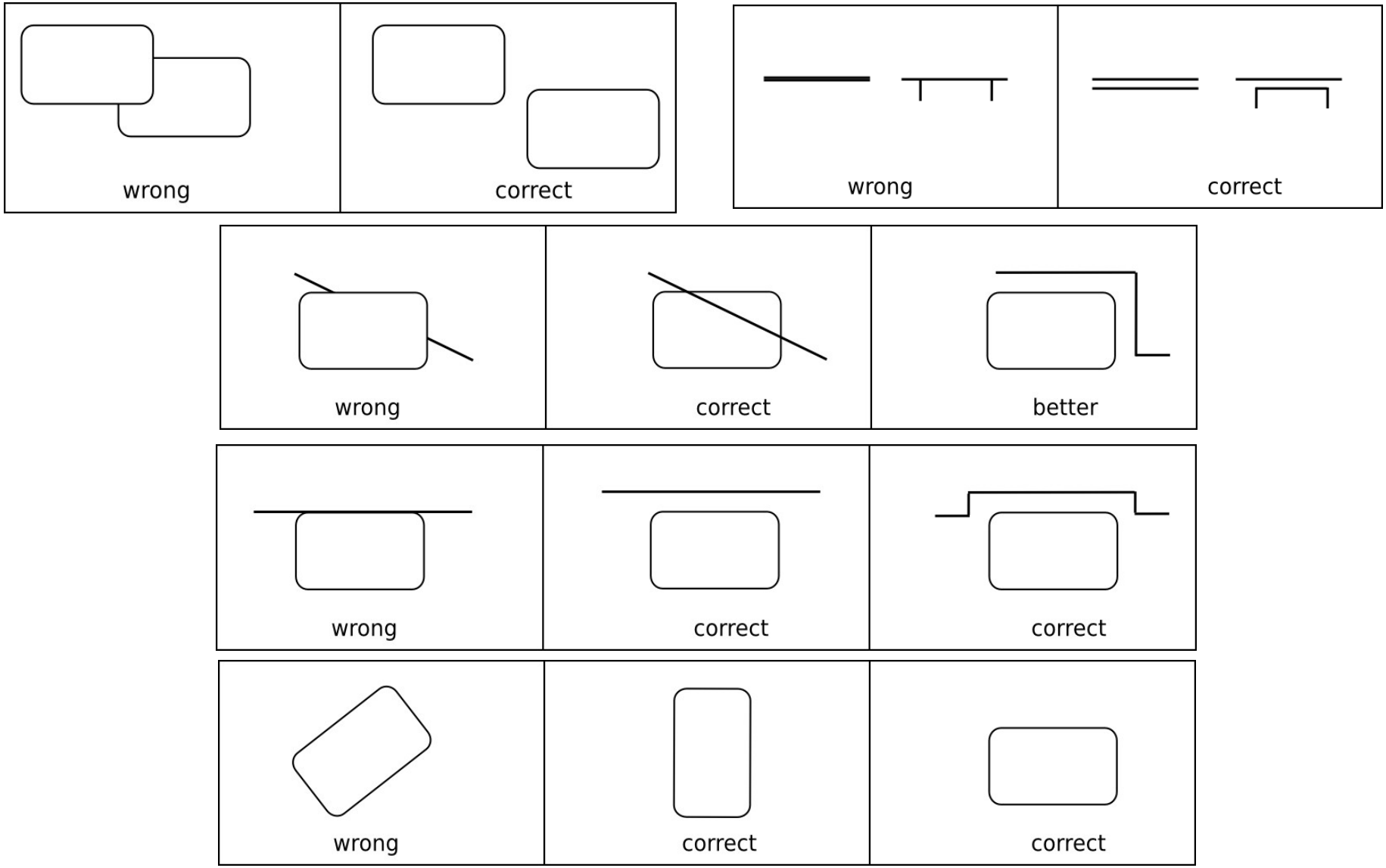

Entity Pool Nodes
LABEL
unspecified entity
LABEL
simple chemical
LABEL
macromolecule
LABEL
nucleic acid feature
LABEL
perturbing agent
source
sink
INFO
varZ
LABEL
INFO
varW
LABEL
LABEL
INFO
varY
varX
complex
Auxiliary Units
pre:label
unit of information
val@var
state variable
LABEL
marker
LABEL
clone marker
N:5
LABEL
N:5
LABEL
multimers
N:2
LABEL
N:2
LABEL
Process Nodes
process
omitted process
?
uncertain process
association
dissociation
LABEL
phenotype
Container Nodes
e:INFO
LABEL
compartment
Reference Nodes
EPN
or CN
A
LABEL
B
C
EPN
or CN
EPN
or CN
submap
LABEL
tag
Connecting arcs
Source
EPN
N
consumption
N
Target
EPN
production
Target
PN
modulation
Target
PN
stimulation
Target
PN
catalysis
Target
PN
inhibition
Target
PN
necessary stimulation
Source
EPN
Logical
operator
logic arc
LABEL
EPN
equivalence arc
Logical Operators
AND
and operator
OR
or operator
NOT
not operator

Syntax definition

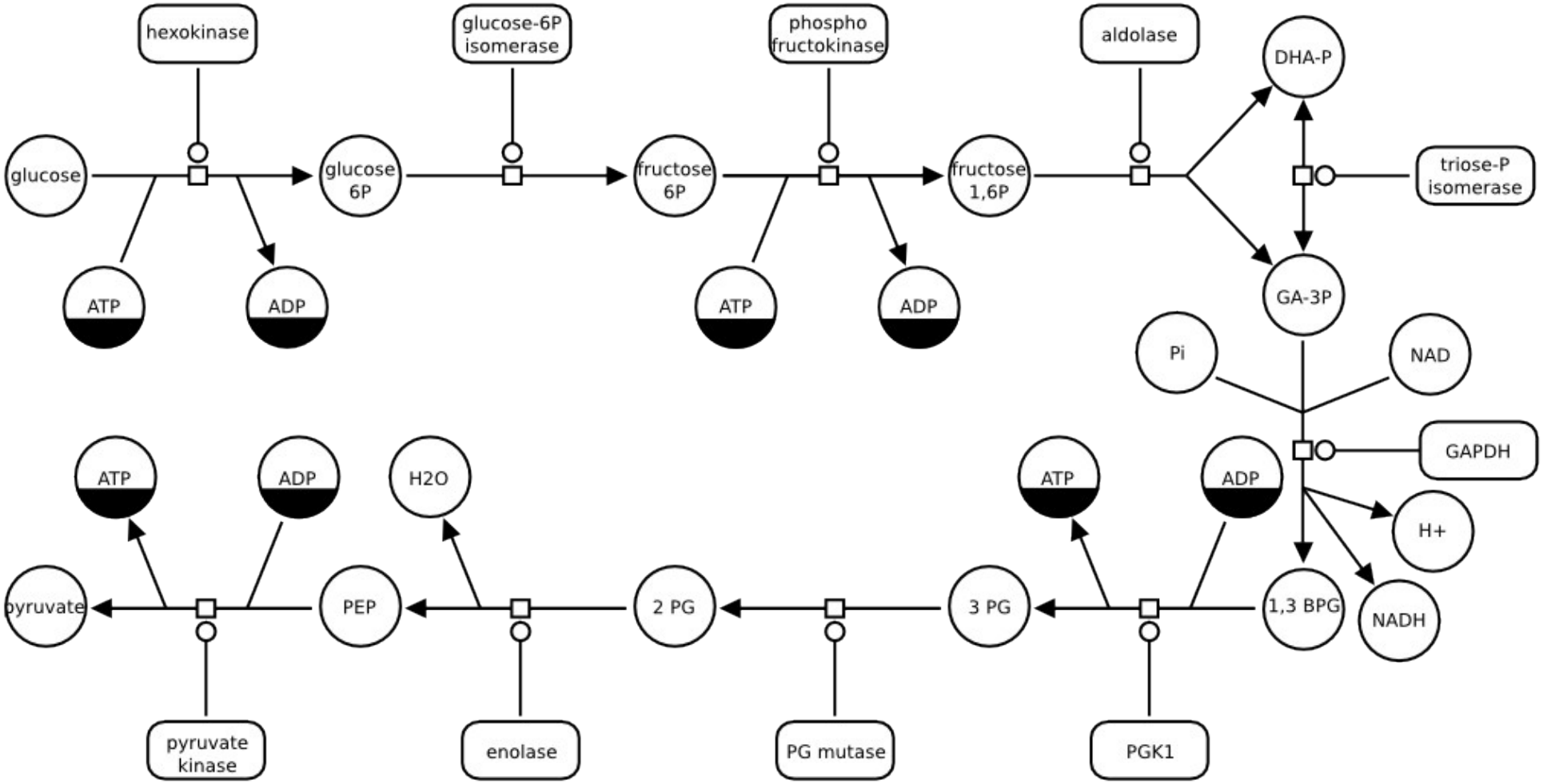
Arc\EPN	*macromolecule*	*simple chemical*	*unspecified entity*	*multimer*	*complex*	*nucleic acid feature*	*tag*	*source/sink*	*perturbation*	*observable*	*submap*
consumption	I	I	I	I	I	I		I			
production	O	O	O	O	O	O		O			
modulation	I	I	I	I	I	I			I	O	
stimulation	I	I	I	I	I	I			I	O	
catalysis	I	I	I	I	I				I	O	
inhibition	I	I	I	I	I	I			I	O	
trigger	I	I	I	I	I	I			I	O	
logic arc	I	I	I	I	I	I					
equivalence arc	I	I	I	I	I	I	O				O

Arc\PN	*transition*	*omitted process*	*uncertain process*	*association*	*dissociation*	*and*	*or*	*not*
consumption	O	O	O	O	O(1)			
production	I	I	I	I(1)	I			
modulation	O	O	O			I(1)	I(1)	I(1)
stimulation	O	O	O			I(1)	I(1)	I(1)
catalysis	O	O	O			I(1)	I(1)	I(1)
inhibition	O	O	O			I(1)	I(1)	I(1)
trigger	O	O	O			I(1)	I(1)	I(1)
logic arc						O	O	O(1)
equivalence arc								

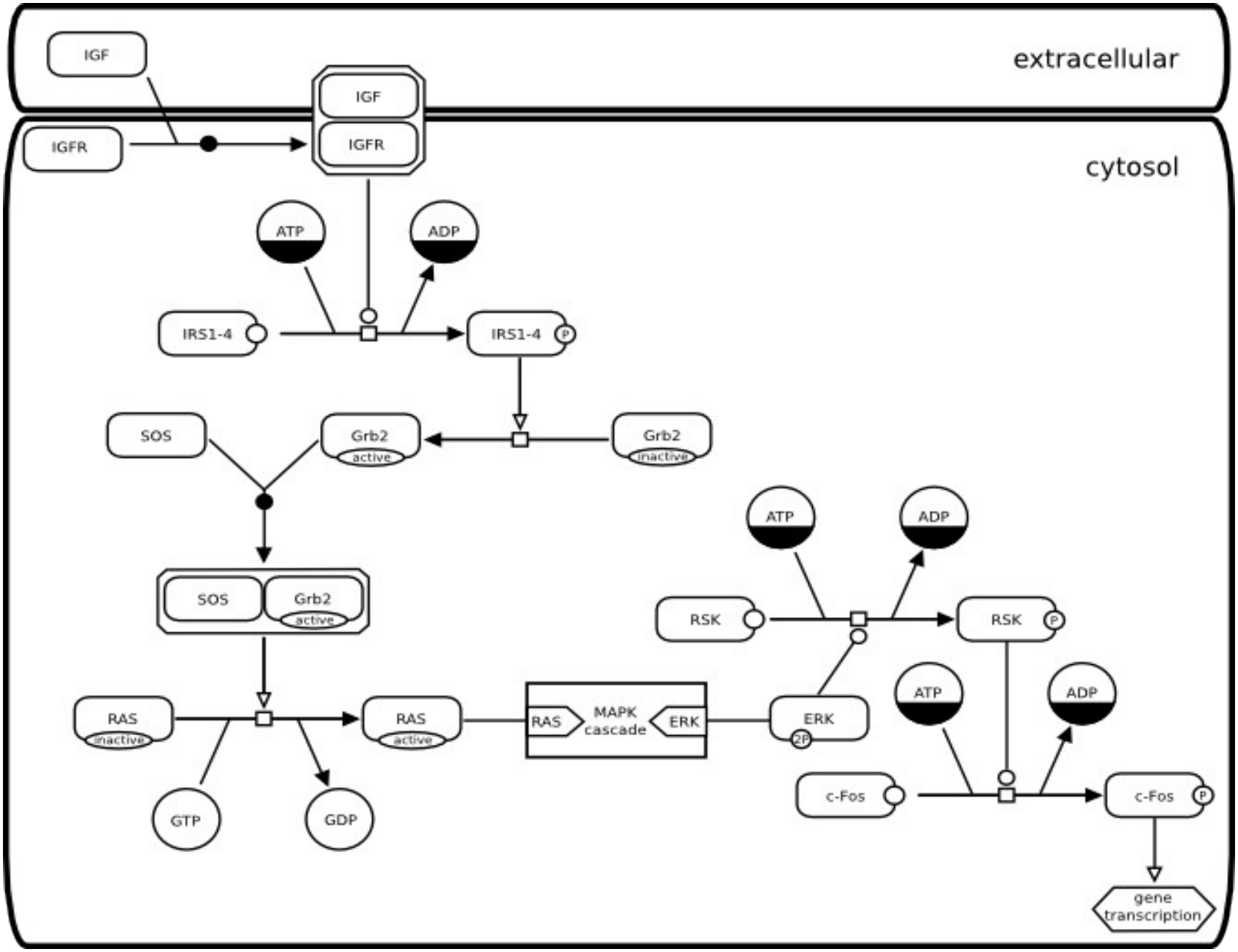
Layout constraints



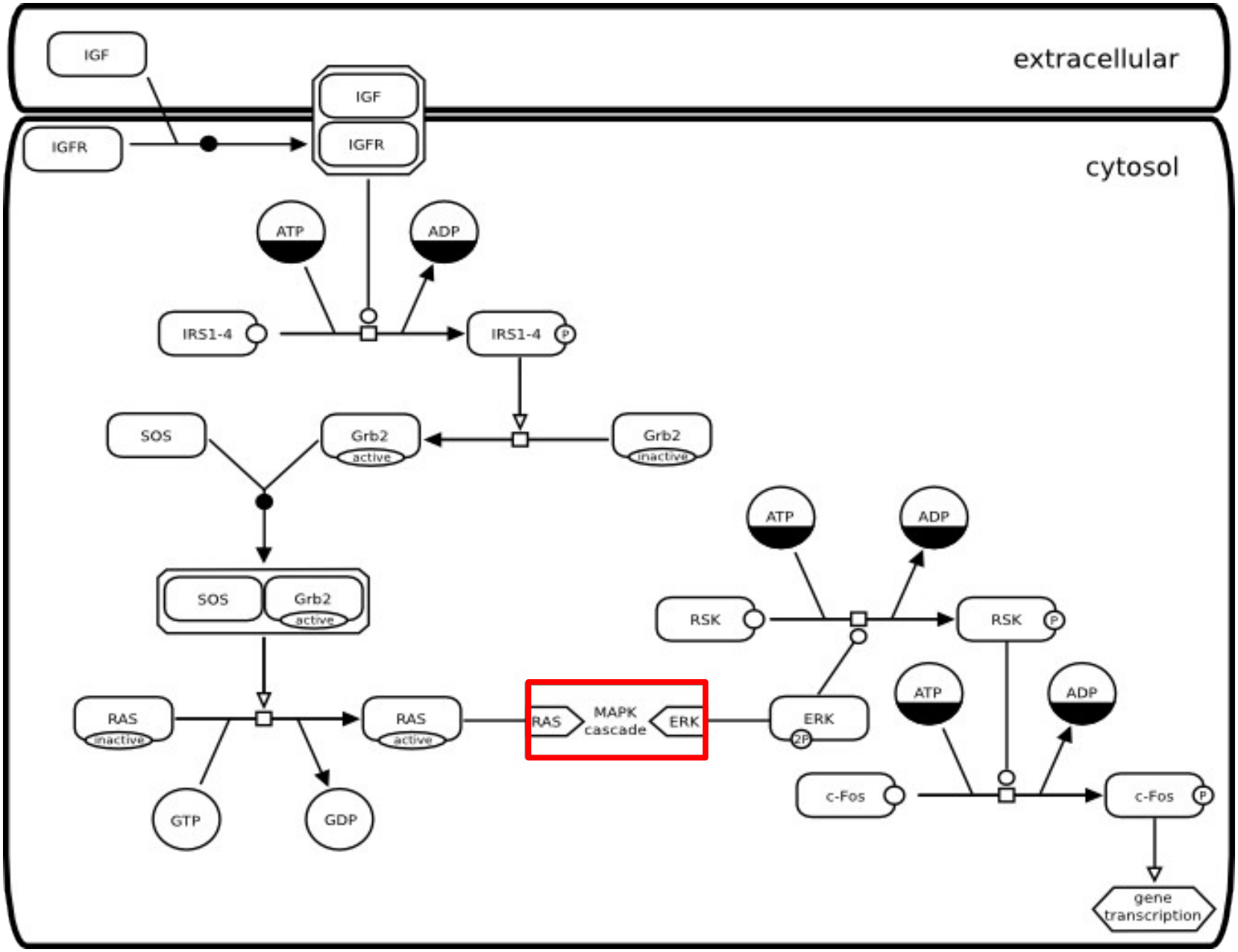
Metabolic network



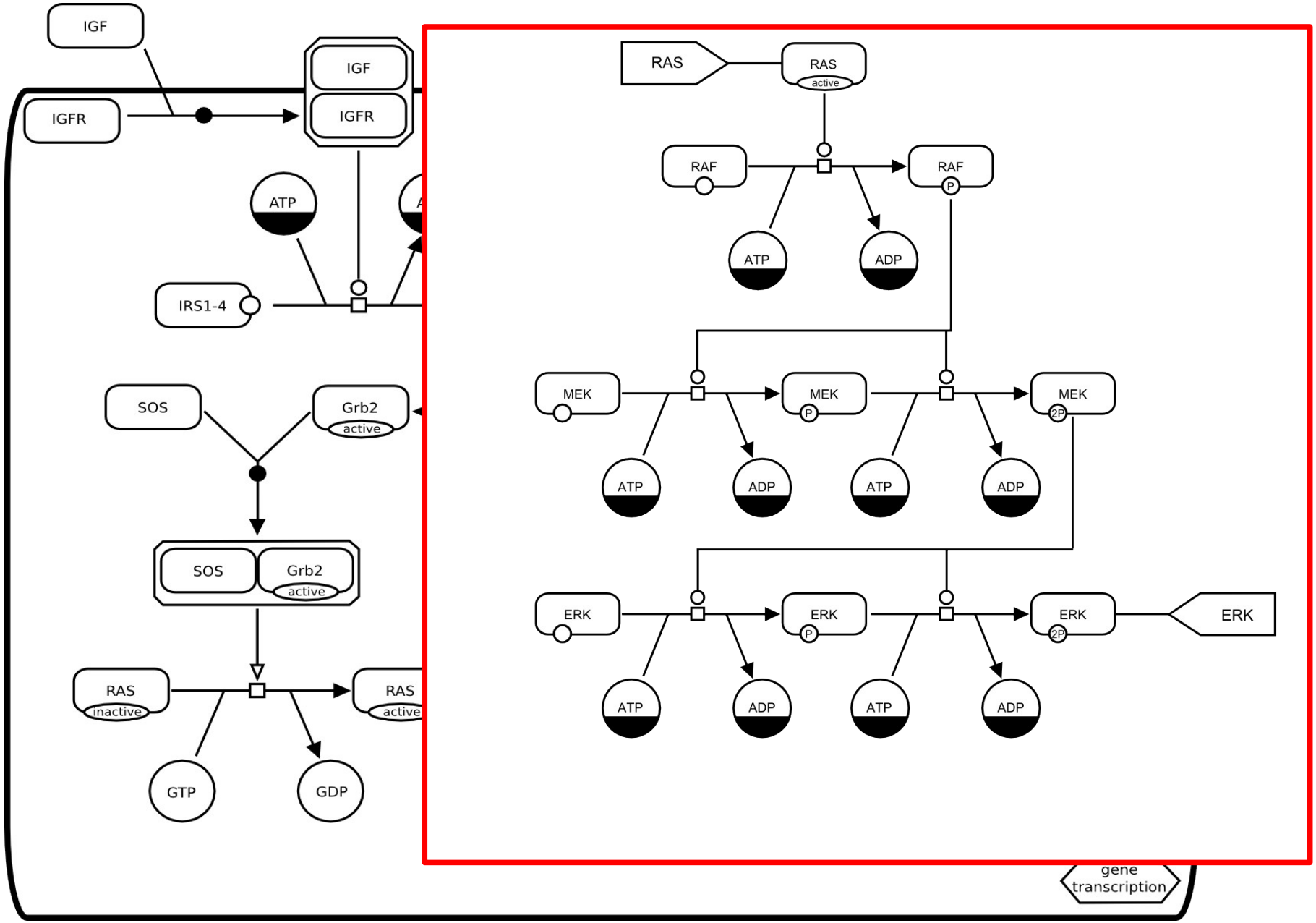
Signalling pathways



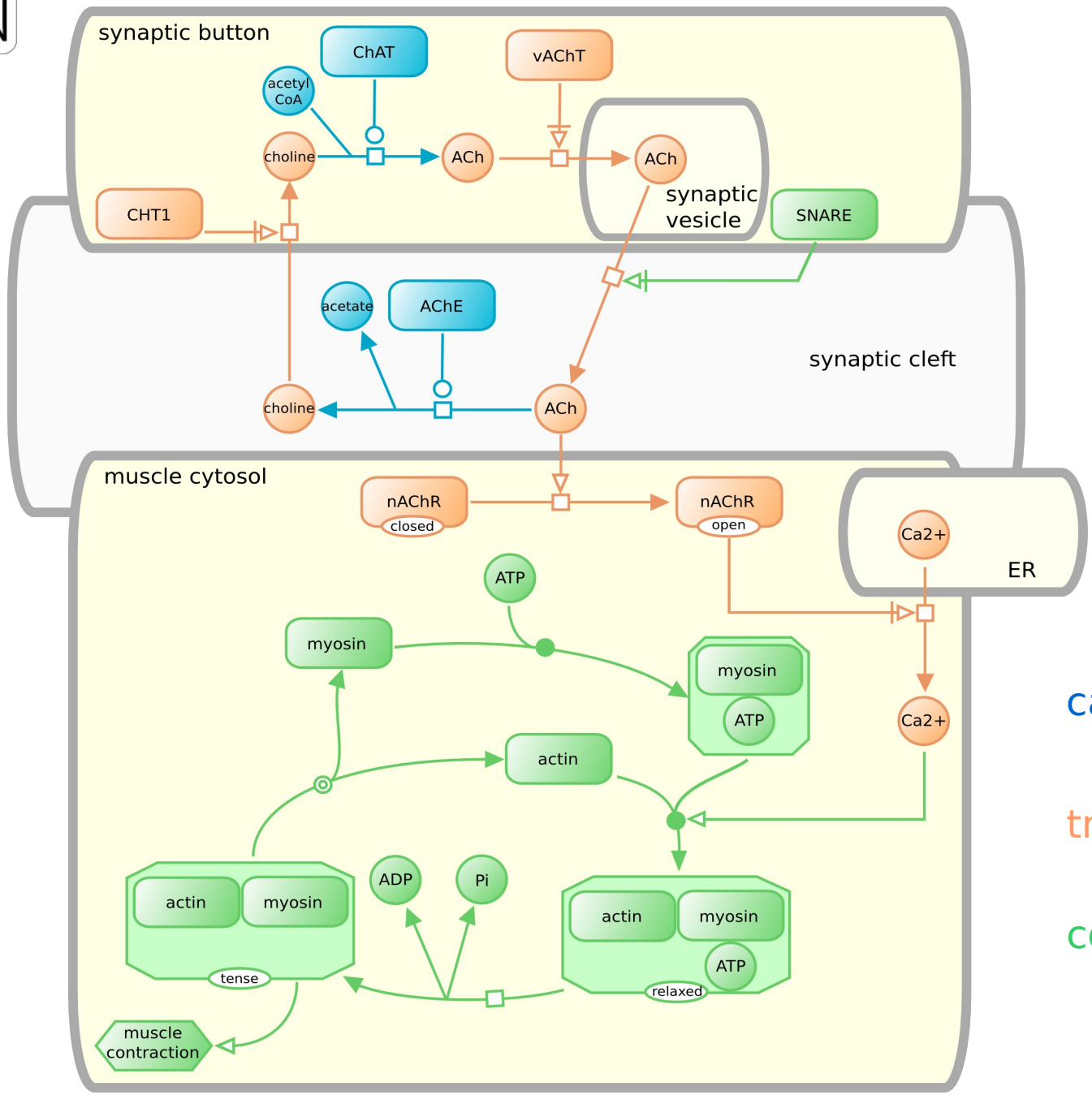
Submaps


extracellular
IGF
IGF
IGFR
IGFR
cytosol
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

Submaps


IGF
IGFR
IGF
IGFR
ATP
IRS1-4
SOS
Grb2
active
SOS
Grb2
active
RAS
inactive
RAS
active
GTP
GDP
RAS
RAS
active
RAF
RAF
P
ATP
ADP
MEK
MEK
P
MEK
2P
ATP
ADP
ATP
ADP
ERK
ERK
P
ERK
2P
ERK
ATP
ADP
ATP
ADP
gene
transcription

multi-cellular processes

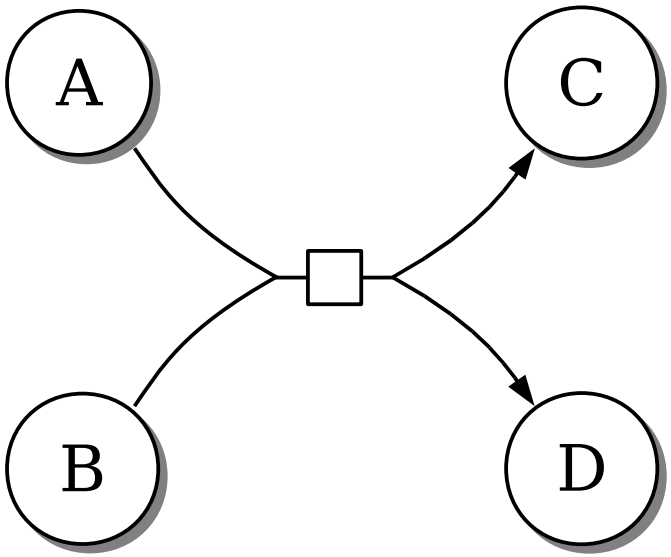


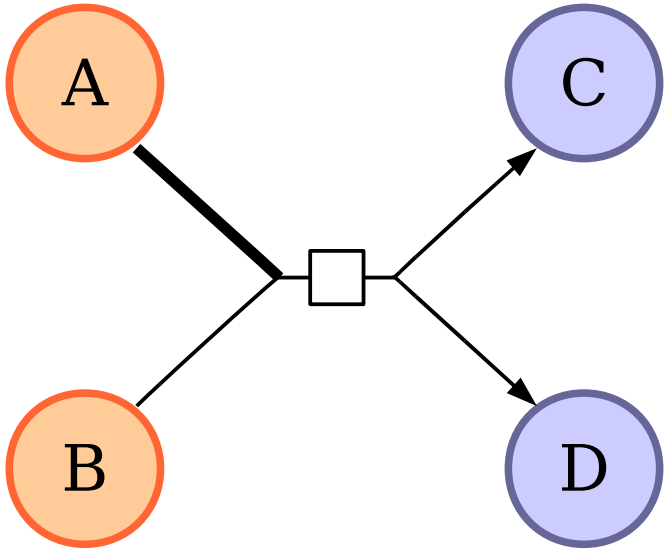
catalytic processes

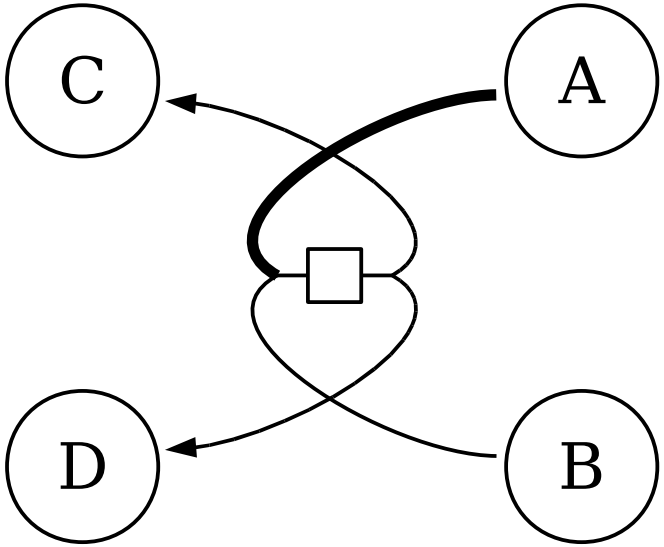
transport processes

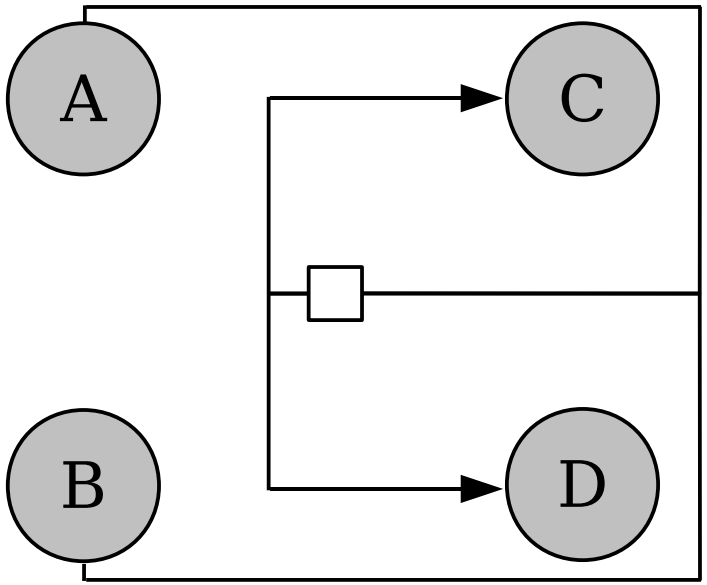
contractile proteins

All those diagrams are identical

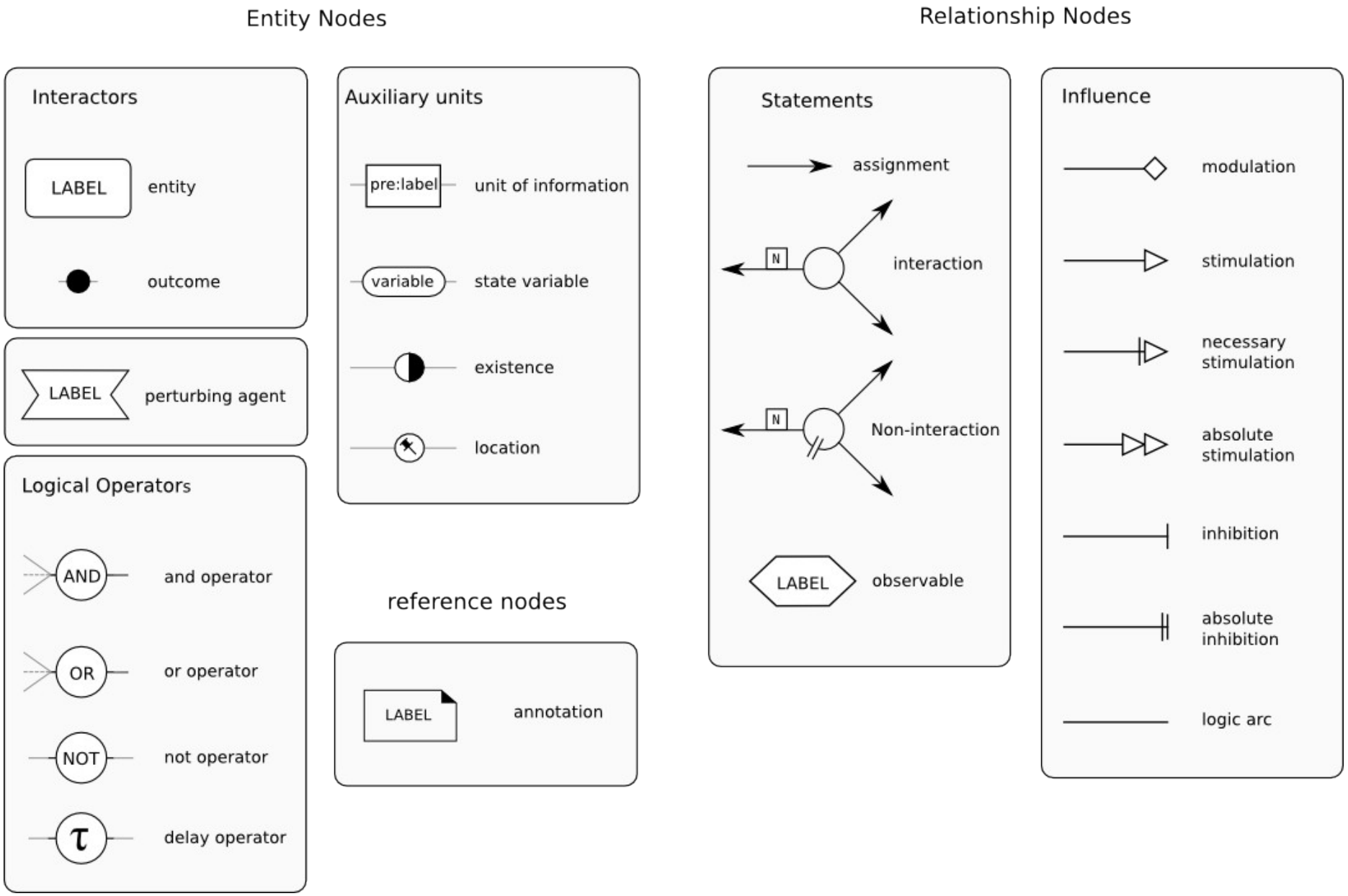




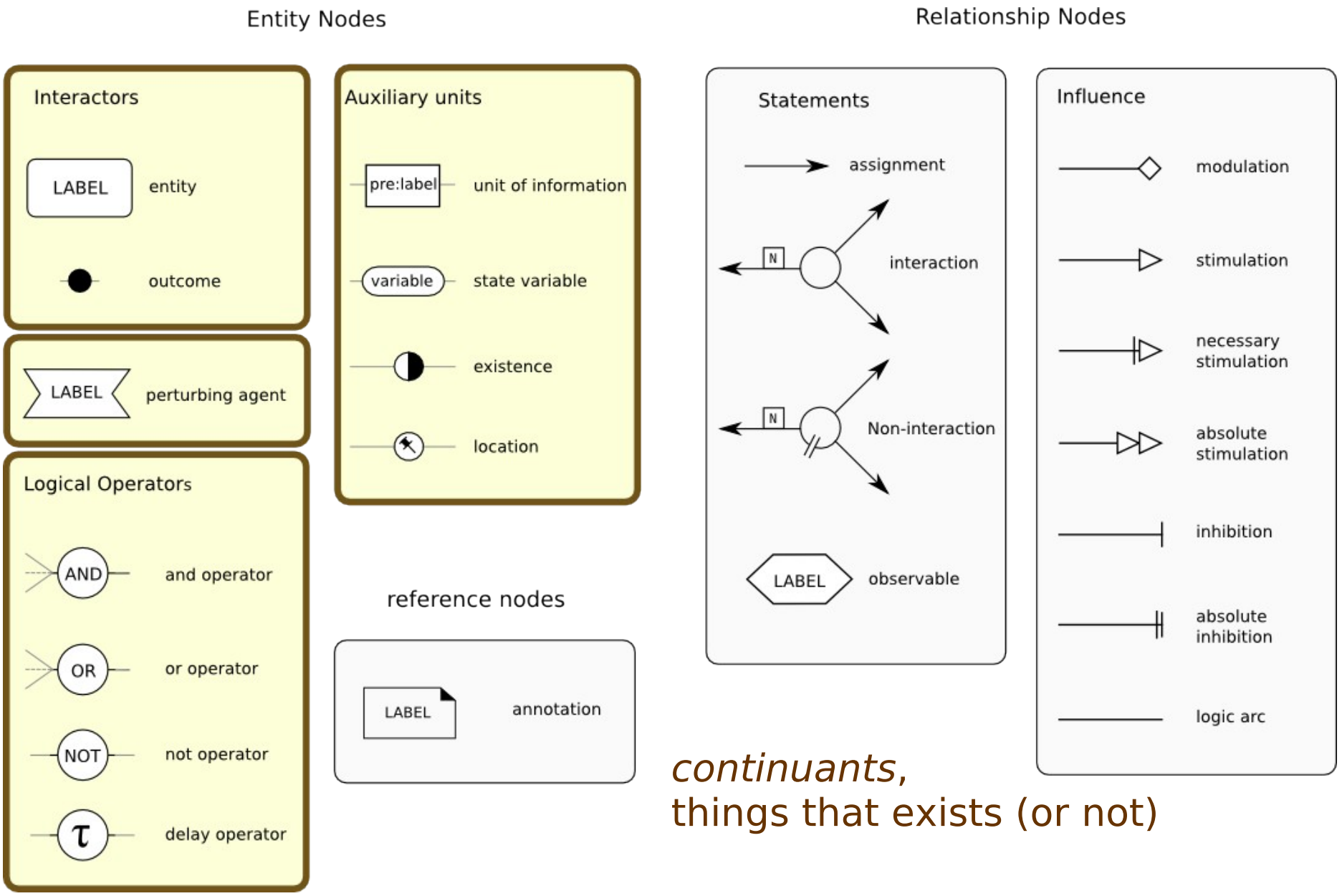




SBGN Entity Relationships L1 reference card

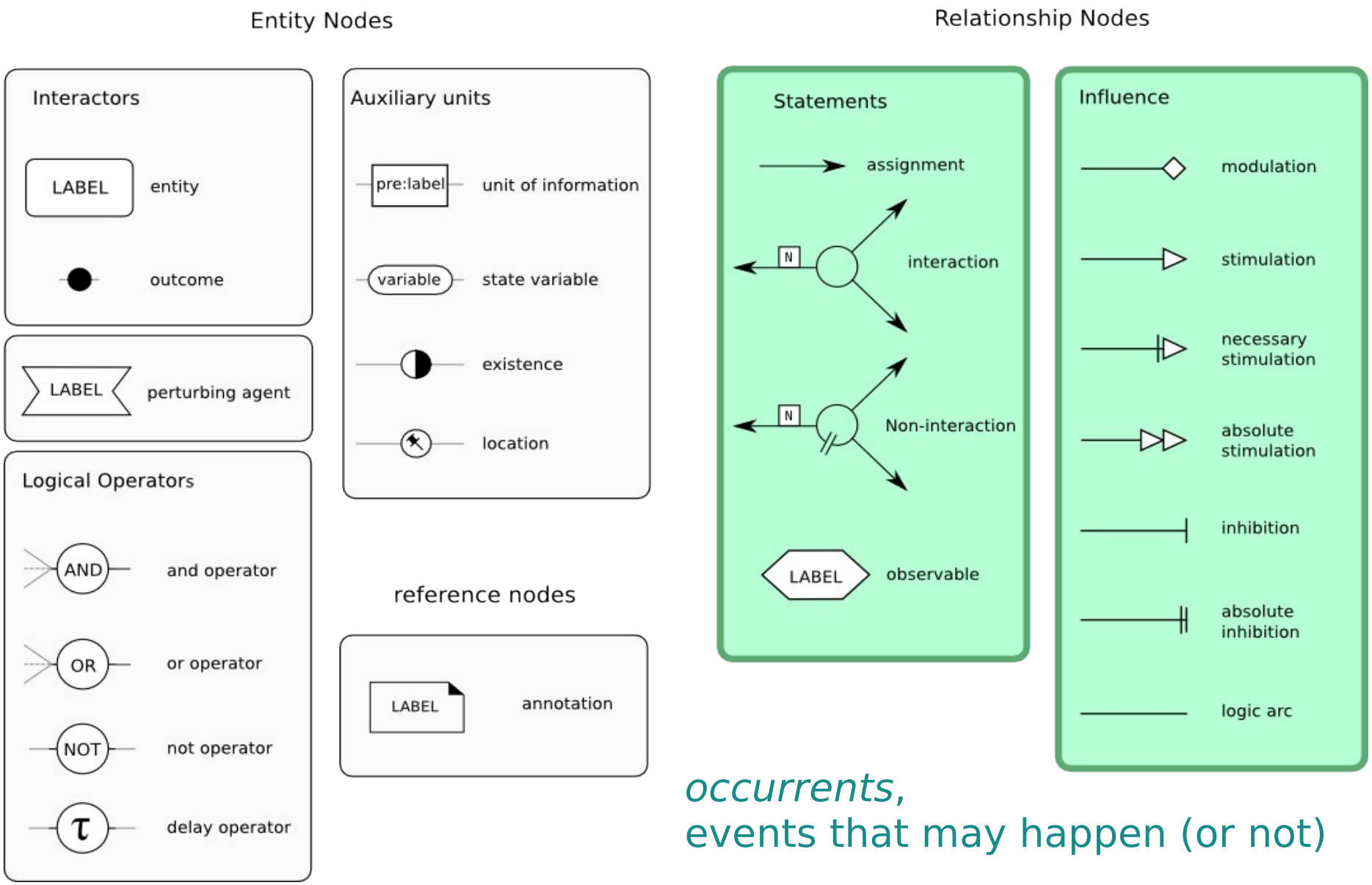


SBGN Entity Relationships L1 reference card



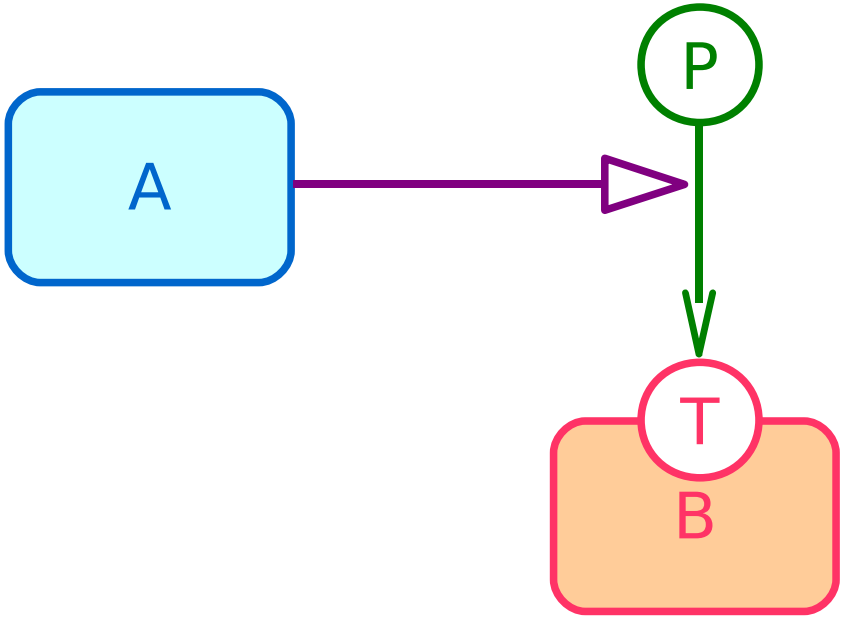
continuants,
things that exists (or not)

SBGN Entity Relationships L1 reference card

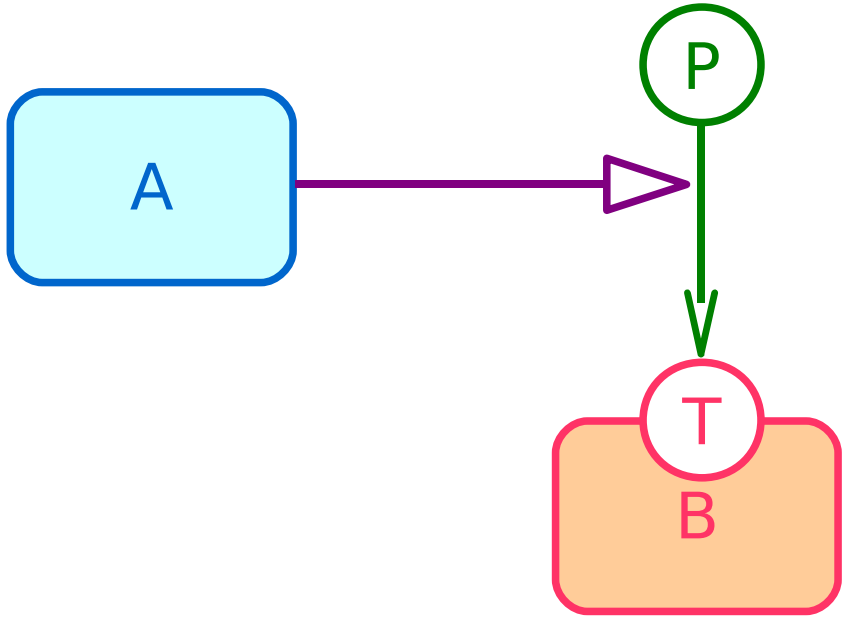


occurrents,
events that may happen (or not)

Entity Relationships can be viewed as rules

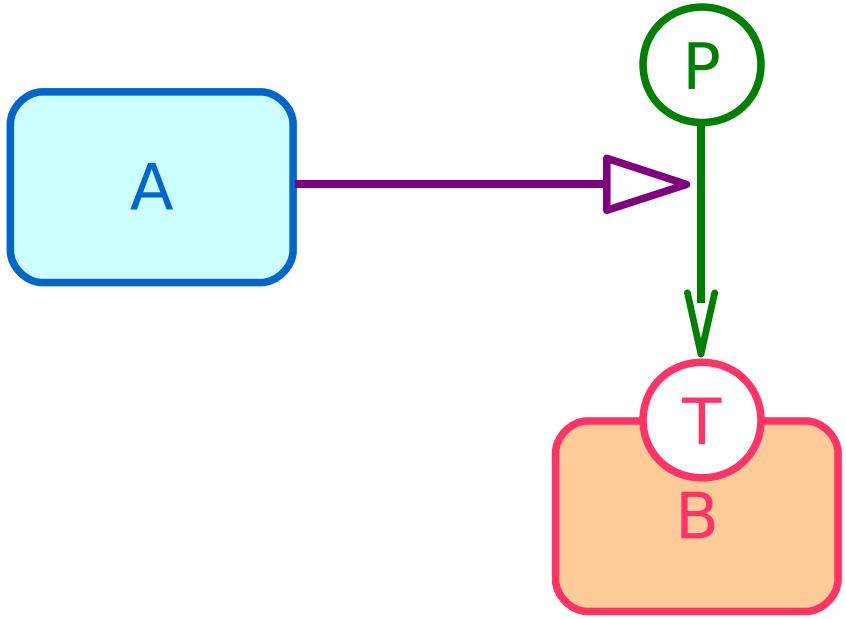


Entity Relationships can be viewed as rules



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

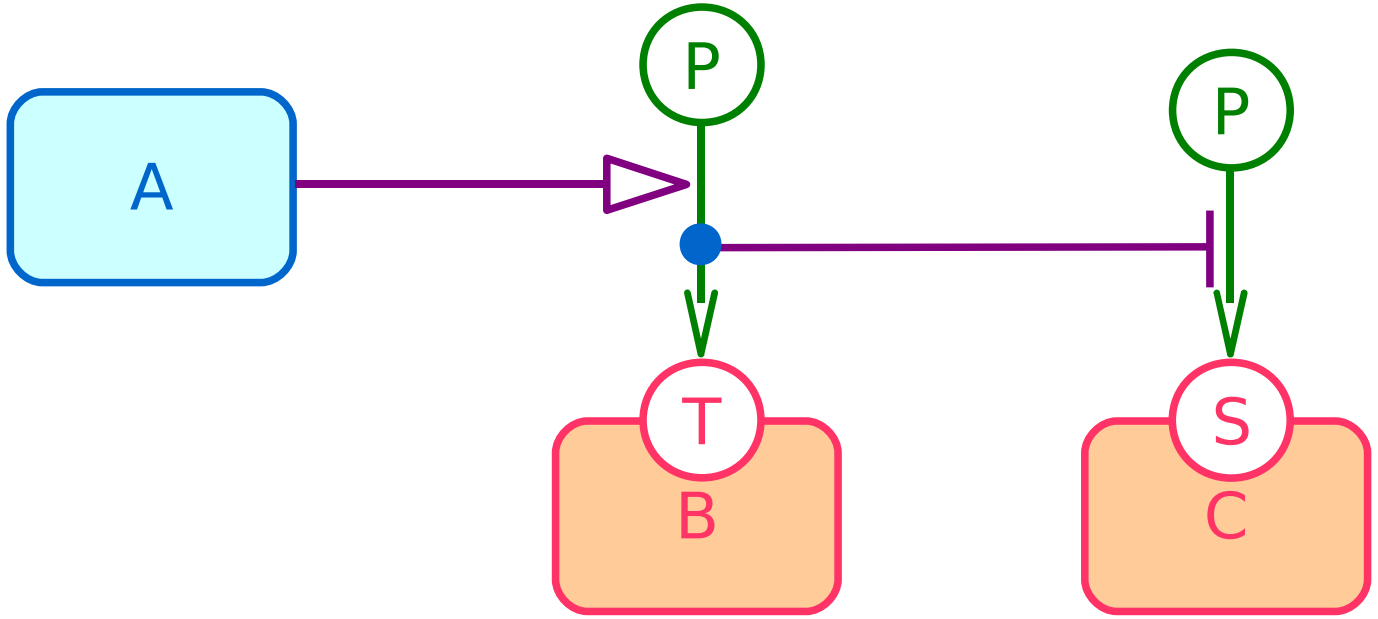
Entity Relationships can be viewed as rules



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)

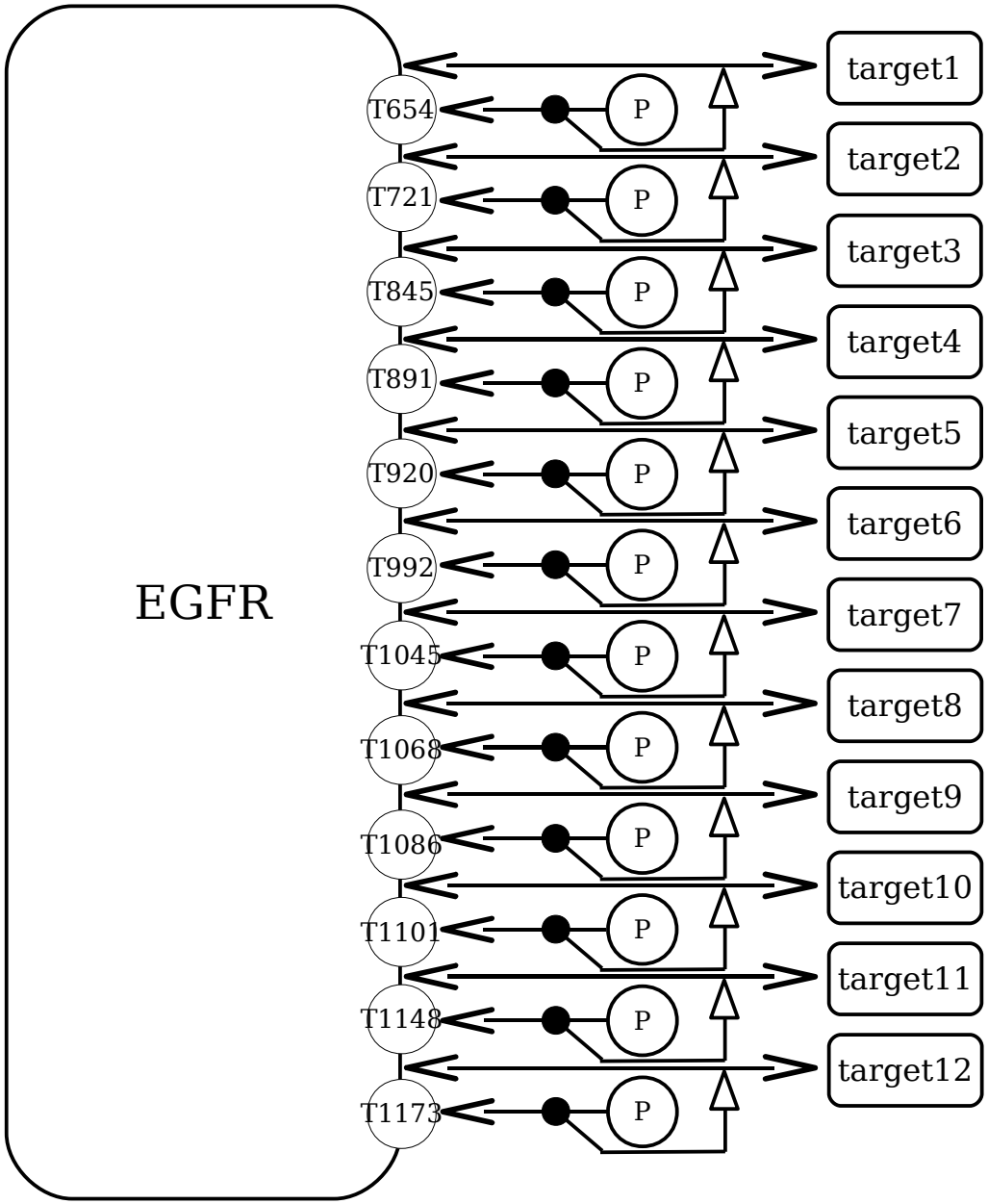
Entity Relationships can be viewed as rules



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

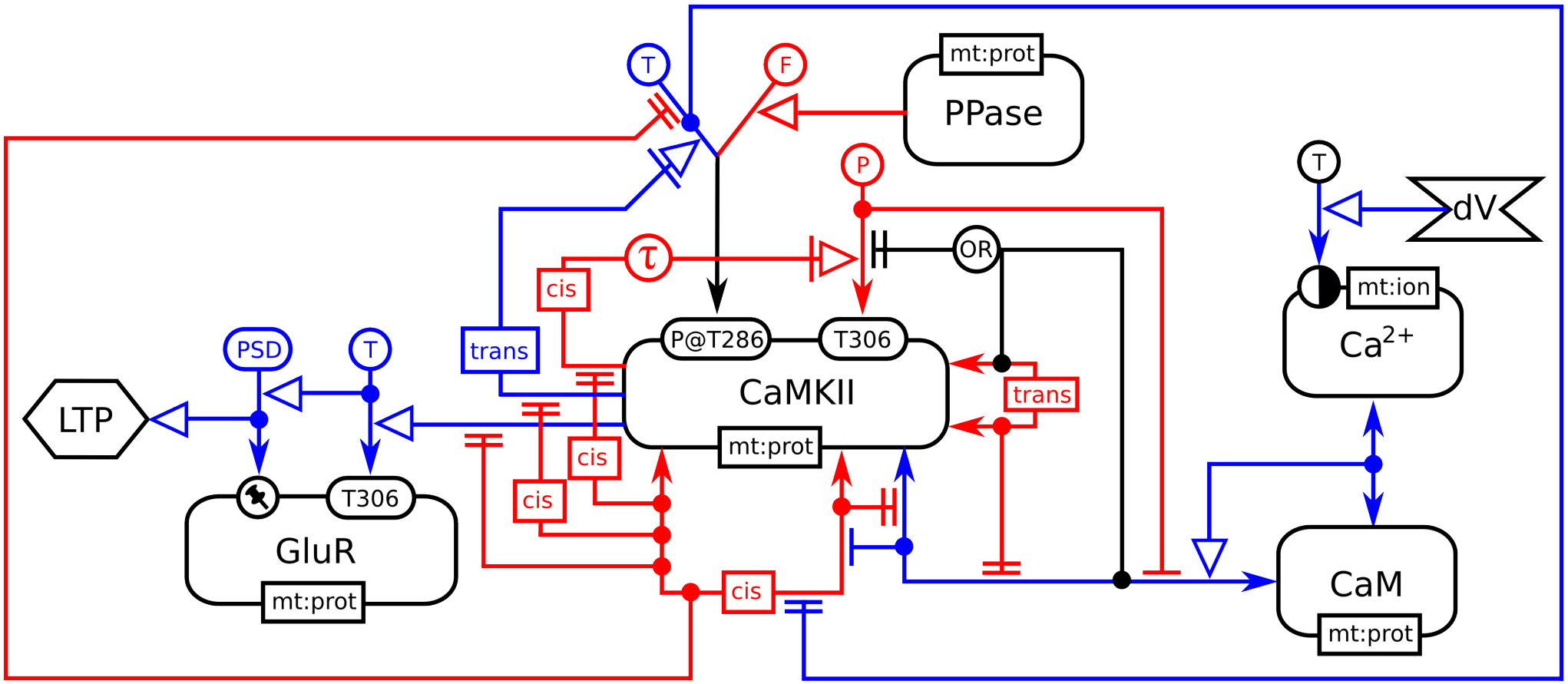
Multistate and combinatorial explosion



Process Diagram:
“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

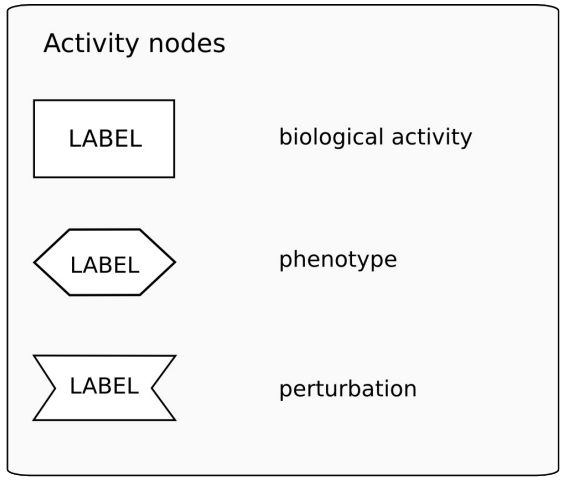
Example of Entity Relationships map

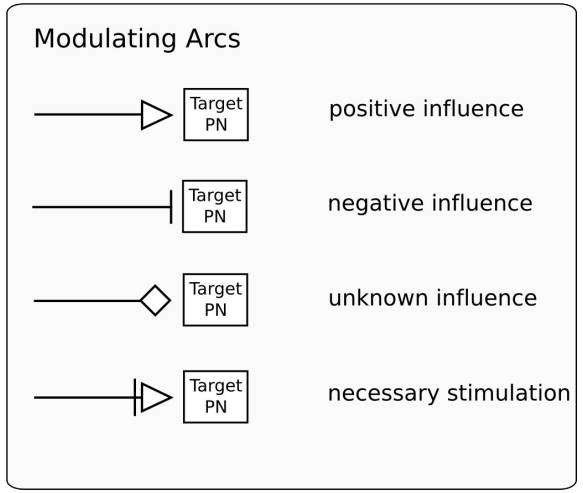


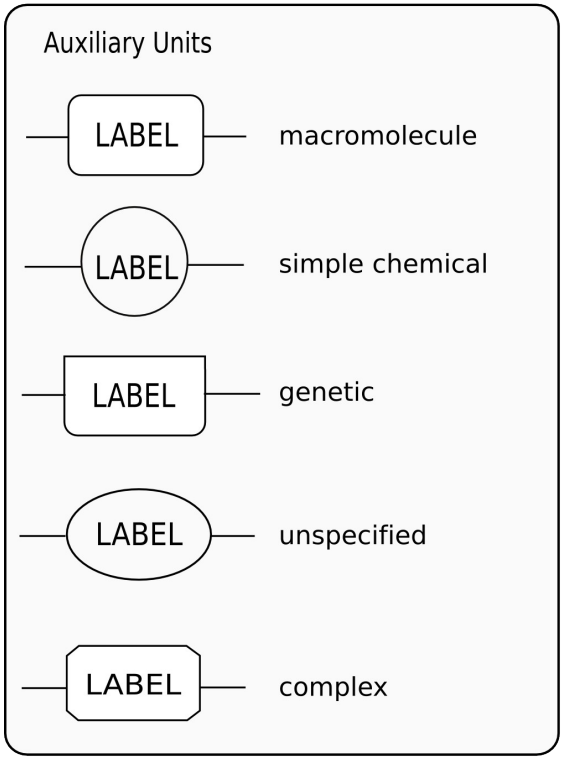
increases synaptic weight

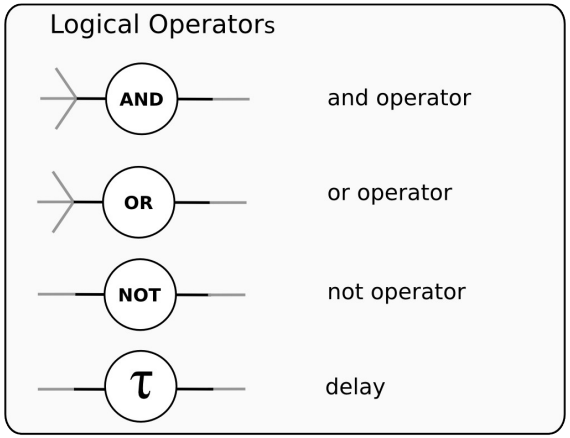
decreases synaptic weight

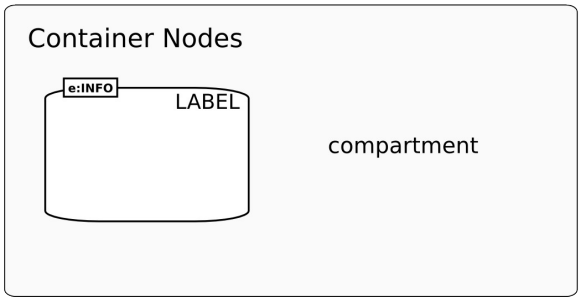
SBGN Activity Flow L1 reference card



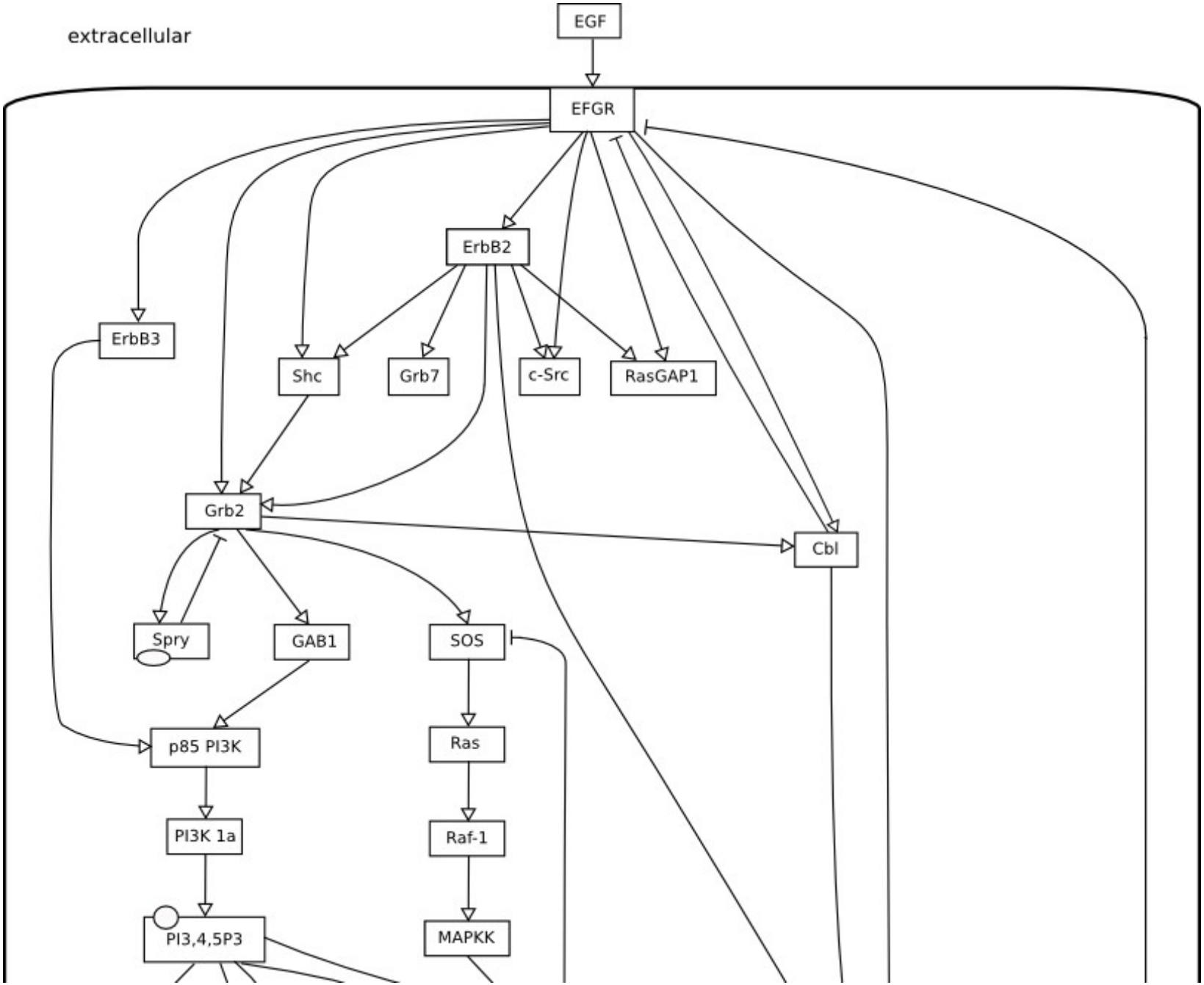


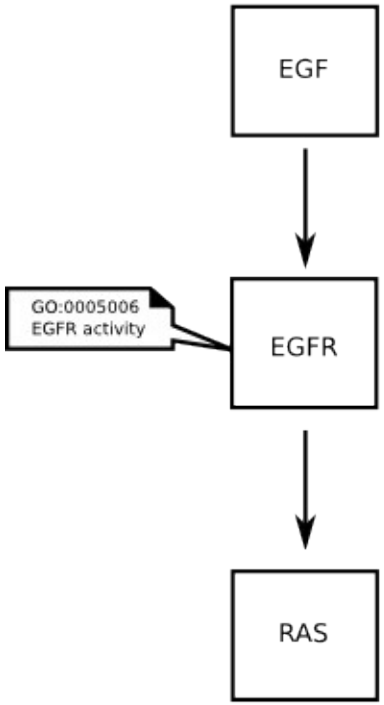


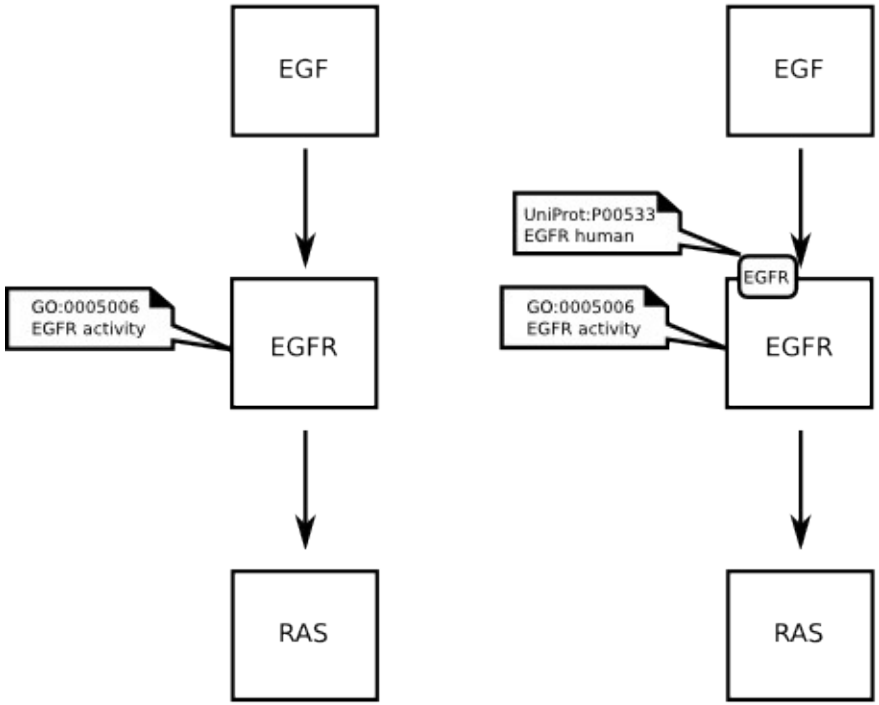


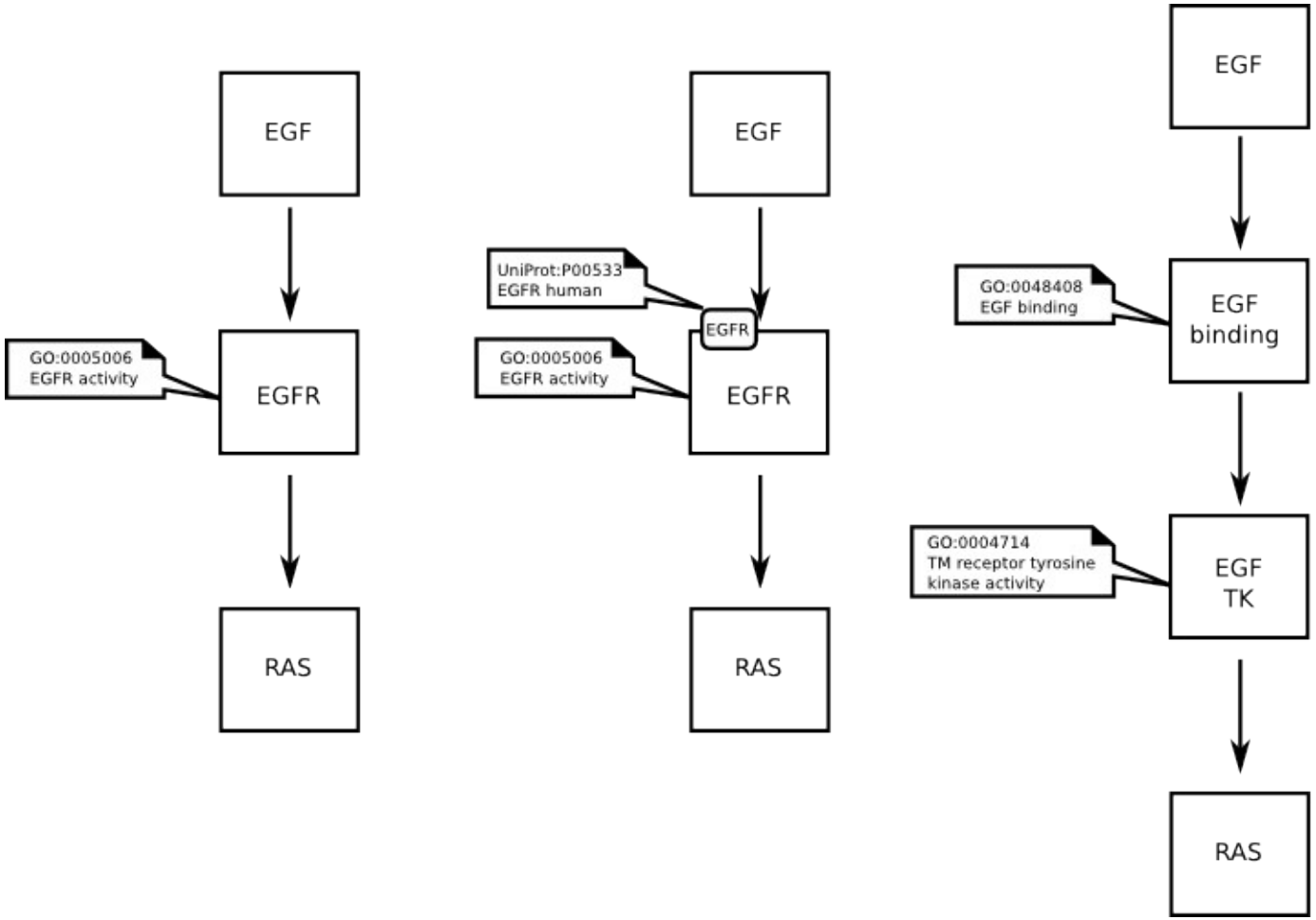


Example of Activity Flow map

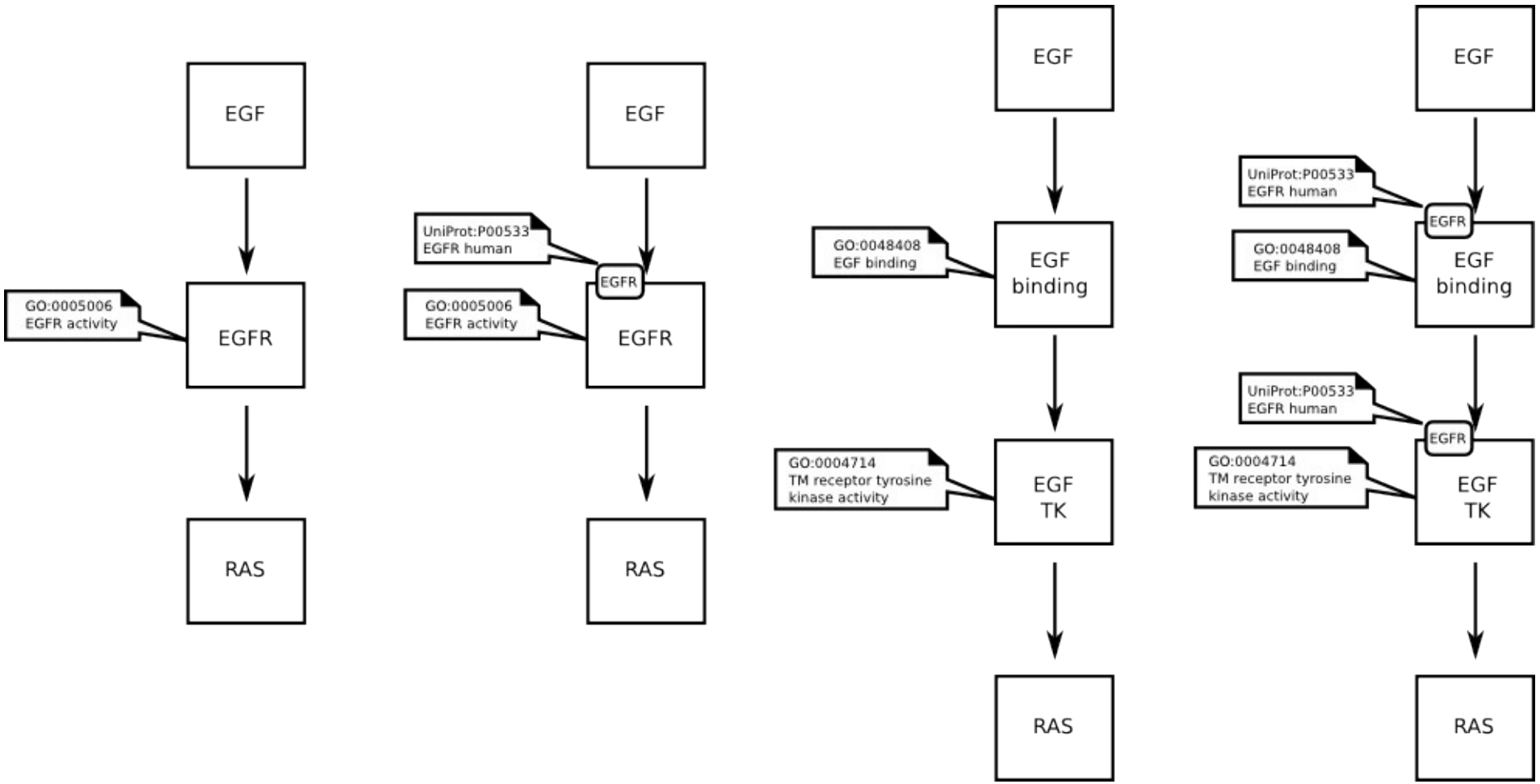



EGF
GO:0005006
EGFR activity
EGFR
RAS


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

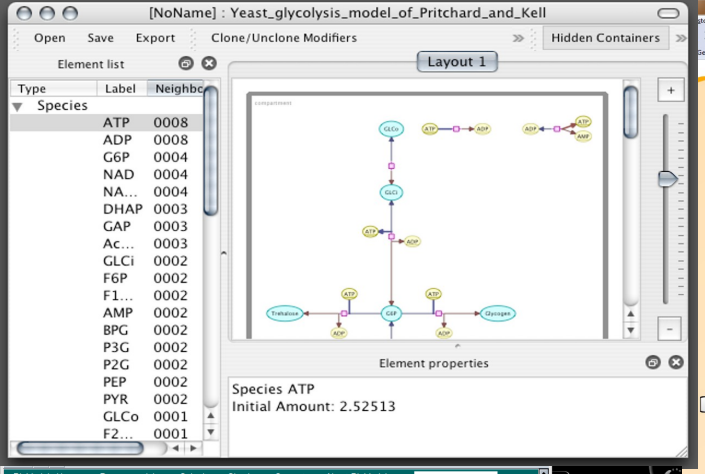

EGF
GO:0005006
EGFR activity
EGFR
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0005006
EGFR activity
EGFR
RAS
EGF
GO:0048408
EGF binding
EGF
binding
GO:0004714
TM receptor tyrosine
kinase activity
EGF
TK
RAS

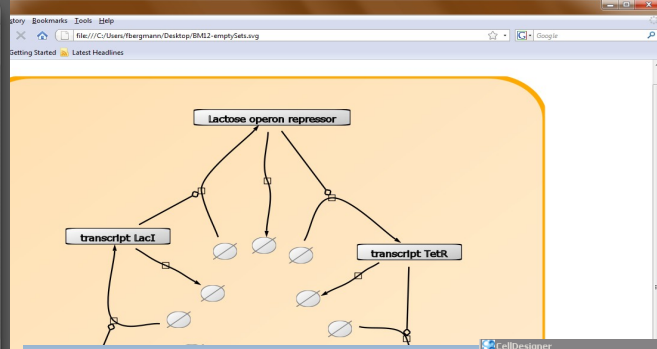
TIMTOWTDI

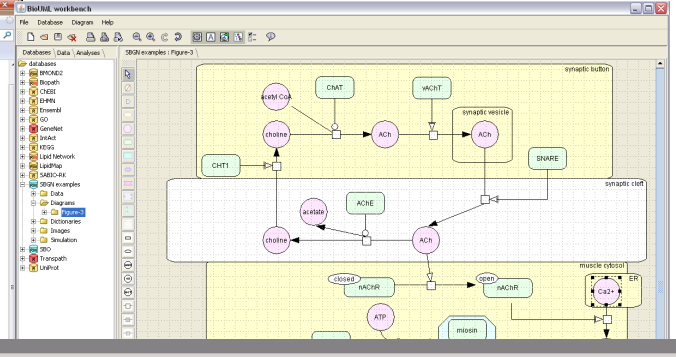

EGF
EGFR
GO:0005006
EGFR activity
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0005006
EGFR activity
EGFR
RAS
EGF
GO:0048408
EGF binding
EGF
binding
GO:0004714
TM receptor tyrosine
kinase activity
EGF
TK
RAS
EGF
UniProt:P00533
EGFR human
EGFR
GO:0048408
EGF binding
EGF
binding
UniProt:P00533
EGFR human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

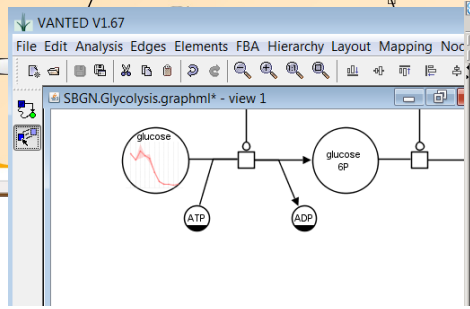
Software support for SBGN PD L1

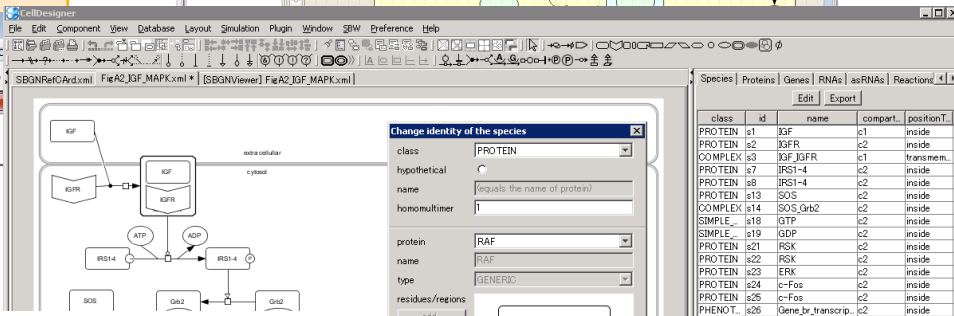
- **Arcadia** (MCISB, Manchester, UK) http://arcadiapathways.sourceforge.net/
- **Athena** (Univ Washington, Seattle, US) http://www.codeplex.com/athena/
- **BioModels DB** (EMBL-EBI, UK) http://www.ebi.ac.uk/biomodels/
- **BioUML** (Inst Systems Biology, Novosibirsk, RU) http://www.biouml.org/
- **CellDesigner** (SBI, Tokyo, JP) http://www.celldesigner.org/
- **EPE** (CISBE, Edinburgh, UK) http://www.bioinformatics.ed.ac.uk/epe/
- **JWS Online** (Stellenbosh University, ZA) http://jjj.biochem.sun.ac.za/
- **NetBuilder** (Univ Hertfords, UK) http://strc.herts.ac.uk/bio/maria/Apostrophe/
- **PANTHER** (SRI international, USA) http://www.pantherdb.org/pathway/
- **Reactome** (EMBL-EBI, UK) http://www.reactome.org/
- **Vanted** (IPK Gatersleben, DE) http://vanted.ipk-gatersleben.de/
- **VISIOweb** (Bilkent Univ, Turkey) http://www.bilkent.edu.tr/~bcbi/pvs.html

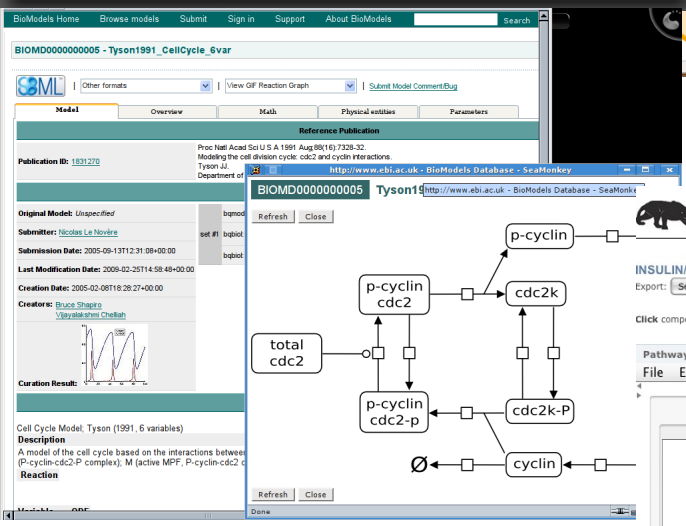

[NoName] : Yeast_glycolysis_model_of_Pritchard_and_Kell
Open
Save
Export
Clone/Unclone Modifiers
Hidden Containers
Element list
Layout 1
Type
Label
Species
ATP 0008
ADP 0008
G6P 0004
NAD 0004
DHAP 0003
GAP 0003
GLCi 0002
F6P 0002
PEP 0002
PYR 0002
GLCo 0001
Element properties
Species ATP
Initial Amount: 2.52513


Lactose operon repressor
transcript LacI
transcript TetR

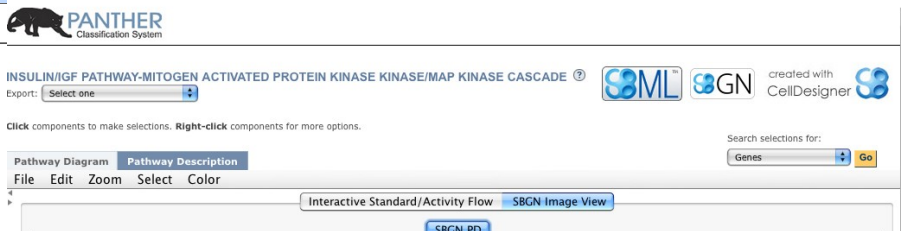

BioUML workbench
synaptic button
synaptic vesicle
synaptic cleft
muscle cytosol

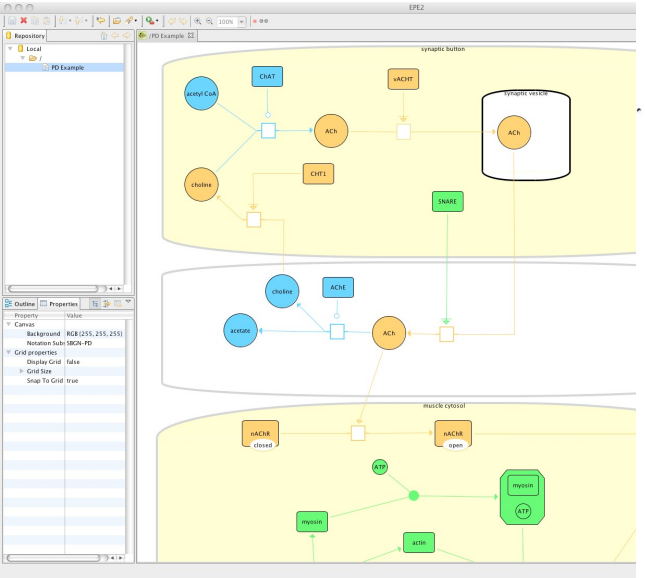

VANTED V1.67
File Edit Analysis Edges Elements FBA Hierarchy Layout Mapping
SBGN.Glycolysis.graphml* - view 1
glucose
glucose 6P
ATP
ADP

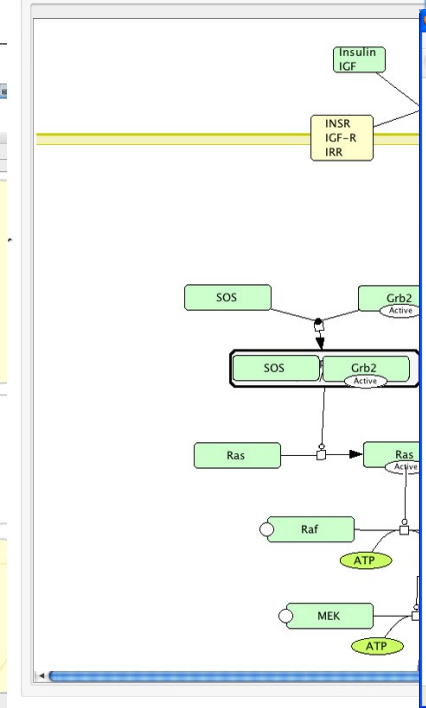

CellDesigner
Change identity of the species
class
PROTEIN
hypothetical
name
homomultimer
protein
RAF
type
GENERIC
residues/regions
modification
state
empty
text input
Apply
Reset
Cancel

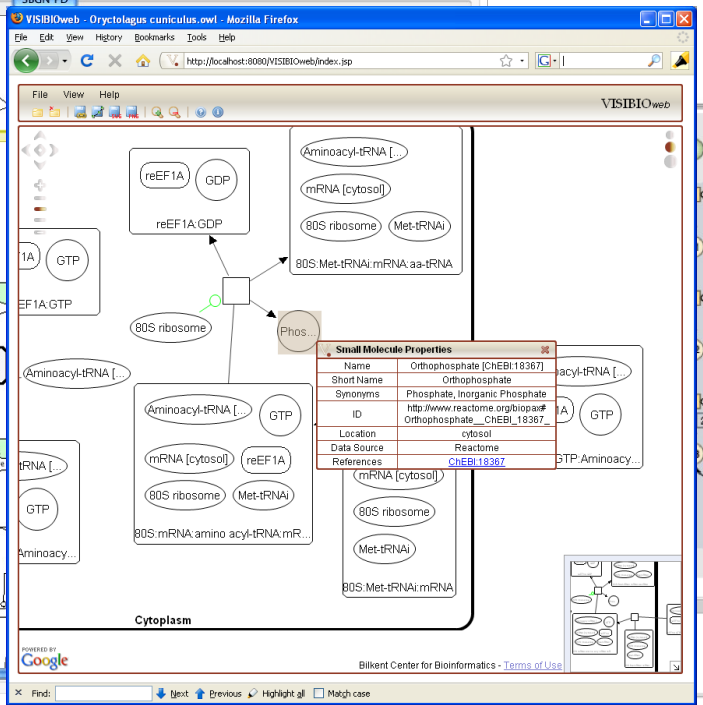

BioModels Home
Browse models
Submit
Sign in
Support
About BioModels
BIOMD0000000005 - Tyson1991_CellCycle_6var
Reference Publication
Cell Cycle Model; Tyson (1991, 6 variables)
Description
Reaction

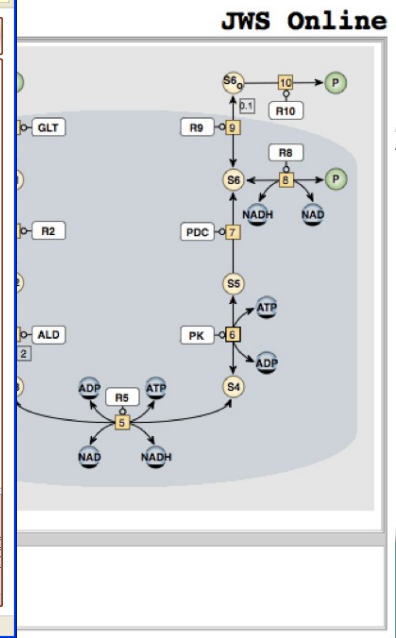
BIOMD0000000005
Refresh
Close
p-cyclin
p-cyclin cdc2
cdc2k
total cdc2
p-cyclin cdc2-p
cdc2k-P
cyclin


PANTHER
Classification System
INSULIN/IGF PATHWAY-MITOGEN ACTIVATED PROTEIN KINASE KINASE/MAP KINASE CASCADE
created with CellDesigner
Pathway Diagram
Pathway Description
File Edit Zoom Select Color
Interactive Standard/Activity Flow
SBGN Image View
SBGN PD


EPE2
synaptic button
synaptic vesicle
muscle cytosol


Insulin
IGF
INSR
IGF-R
IRR
SOS
Grb2
Ras
Raf
MEK
ATP


VISIBIOweb - Oryctolagus cuniculus.owl - Mozilla Firefox
VISIBIOweb
reEF1A
GDP
reEF1A:GDP
80S ribosome
Met-tRNAi
80S:Met-tRNAi:mRNA:aa-tRNA
Small Molecule Properties
Name
Orthophosphate [ChEBI:18367]
Short Name
Orthophosphate
Synonyms
Phosphate, Inorganic Phosphate
ID
Location
cytosol
Data Source
Reactome
References
ChEBI:18367
Cytoplasm
Bilkent Center for Bioinformatics - Terms of Use


JWS Online
GLT
R9
R10
R8
R2
PDC
NADH
NAD
ALD
PK
ADP
ATP
R5
S4

Status of Specifications

- SBGN Process Diagrams
  - Level 1 Version 1.0 release on August 23rd 2008
  - Level 1 Version 1.1 to be released over summer
  - Level 1 Version 2 to be finalised over summer
- SBGN Entity Relationships
  - Level 1 Version 1.0 to be released over summer
- SBGN Activity Flow
  - Level 1 Version 1.0 to be released over summer

Future SBGN meetings

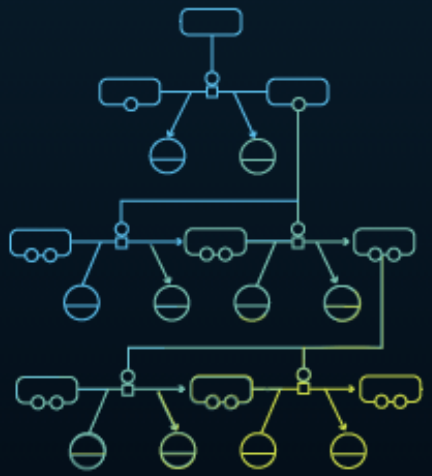
- 5th SBGN forum
  - 02-03 September 2009,
  - San Francisco
  - “Satellite” of ICSB 2009
- 4rd SBGN hackathon (SBGN 5.5)
  - 21-23 April 2010
  - Wittenberg
- 6th SBGN forum (provisional)
  - October 2010
  - Edinburgh
  - Satellite of ICSB 2010
- 5rd SBGN hackathon (SBGN 6.5)
  - Spring 2011, Bethesda, USA

SBGN Systems Biology Graphical Notation

http://www.sbgn.org/

Home | Documents | Forums | Community | Events | About

Google Site Search..



A Visual Notation for Network Diagrams in Biology

SBGN.org is 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 diagrams of biochemical and cellular processes studied in systems biology.

Standardizing the diagrammatic notation is crucial for more efficient and accurate transmission of biological knowledge between different research communities in the expanding field of systems biology. Notations traditionally used by researchers and software have been informal, idiosyncratic and highly variable. Until SBGN, there has been no standard agreed-upon convention defining precisely how to draw biochemical interaction diagrams in a regular and systematic way that helps readers interpret them consistently and unambiguously.

SBGN defines a comprehensive set of symbols with precise semantics, together with detailed syntactic rules defining their use and how diagrams are to be interpreted. By standardizing the visual notation, SBGN can serve as a bridge between different communities in research, education, publishing, and more. The real payoff will come when researchers are as familiar with the notation as electronics engineers are familiar with the notation of circuit schematics. If researchers are saved the time and effort required to familiarize themselves with different notations, they can spend more time thinking about the biology being depicted.

On this site, you can browse some example diagrams to get a feeling for SBGN, read the SBGN specification documents, join online discussions, see current working documents in the SBGN wiki, and much more.

SBGN is the work of many people. It would not have been possible without the generous support of multiple organizations over the years, for which we are very thankful.

SBGN News

(23 Aug.'08) The first **SBGN Process Diagrams Level 1** specification is out! Download the specification and tell us what you think!

(Locked) | Watch page | Page History | Upload file | File list | Page list | My watchlist | Preferences | Help | Log out

Acknowledgements

SBGN editors

- Nicolas Le Novère (UK)
- Stuart Moodie (UK)
- Anatoly Sorokin (UK)
- Michael Hucka (US)
- Falk Schreiber (DE)
- Huaiyu Mi (US)

Special contributors

Emek Demir, Katja Wegner, Mirit Aladjem, Sarala Wimalaratne, Frank Bergman, Ralph Gauges, Peter Ghazal, Kawaji Hideya, Lu Li, Yukiko Matsuoka, Alice Villéger, Sarah Boyd, Laurence Calzone, Melanie Courtot, Ugur Dogrusoz, Akira Funahashi, Sohyoung Kim, Fedor Kolpakov, Augustin Luna, Sven Sahle, Douglas Kell, Kurt Kohn, Hiroaki Kitano

The whole community participating to SBGN meetings, and mailing-lists

NEDO bbsrc biotechnology and biological sciences research council NIGMS EMBL