Systems Biology Markup Language (SBML)

Jannis Uhlendorf

Motivation

Typical representations of biological models:

$$\frac{d\alpha_{out}}{dt} = -\Psi\, S\, J(\alpha_{out}, \alpha_{in})$$
$$\frac{d\alpha_{in}}{dt} = S\, J(\alpha_{out}, \alpha_{in}) - k_{cat\,A}\, A\, \frac{\alpha_{in}}{K_{M\,A} + \alpha_{in}}$$
$$\frac{d\beta}{dt} = k_{cat\,A}\, A\, \frac{\alpha_{in}}{K_{M\,A} + \alpha_{in}} - k_{cat\,X}\, B\, \frac{\beta}{K_{M\,B} + \beta}$$
$$\frac{d\gamma}{dt} = k_{cat\,B}\, B\, \frac{\beta}{K_{M\,B} + \beta} - k_{cat\,D}\, D\, \frac{\gamma}{K_{M\,D} + \gamma}$$
$$\frac{d\delta}{dt} = k_{cat\,D}\, D\, \frac{\gamma}{K_{M\,D} + \gamma} - k_{cat\,E}\, E\, \frac{\delta}{K_{M\,E} + \delta}$$
$$\frac{d\varepsilon}{dt} = k_{cat\,E}\, E\, \frac{\delta}{K_{M\,E} + \delta} - 1.5789 * 10^{-10};\ \varepsilon > 0$$

$$\frac{d\Psi}{dt} = \frac{d\Psi}{dt} = 5 * 10^{6}\ cells/_{mL}\ \ given\ that\ glycolate\ is\ present$$

```
function dy = nop450(t,y)

dy=zeros(7,1);

dy(1)= -y(7)*(6*10^(-12))*0.0463067*(y(1)-y(2));
dy(2)= ((6*10^(-12))*0.0463067*(y(1)-y(2)))-3.3*25.55*(y(2)/(0.53+y(2)));
dy(3)= 3.3*25.55*(y(2)/(0.53+y(2)))-0.0871*25.55*(y(3)/(0.94+y(3)));
dy(4)= .0871*25.55*(y(3)/(0.94+y(3)))- 0.6*25.55*(y(4)/(0.16+y(4)));
dy(5)= 0.6*25.55*(y(4)/(0.16+y(4))) - 25.4*25.55*(y(5)/(20+y(5)));

   if y(6) >  2*10^(-10)
      dy(6)= 25.4*25.55*(y(5)/(20+y(5))) -1.5789*10^(-10)
   else
       dy(6) = 25.4*25.55*(y(5)/(20+y(5)))
   end

   if y(6) > 0.0005
        if y(7) > 1.6*10^11
            dy(7)=0
        else
            dy(7) = 5*10^6
        end
   else
        dy(7) = 0
   end

end
```

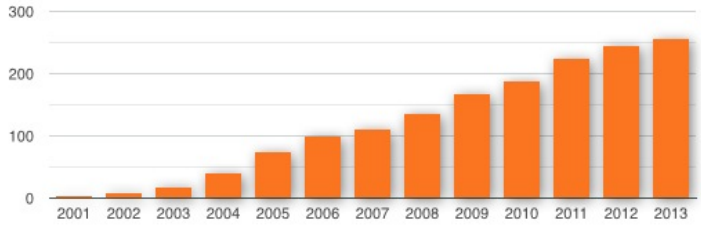
- Problems:
  - How to exchange models between different software tools?
  - How to share models with other researchers?

source:http://2013.igem.org/Team:SydneyUni_Australia/Modelling_Output

- What is SBML?
  - machine readable format for computational models in biology
  - aimed representing biochemical reactions
  - can also be used for other processes (e.g. population models)
  - based on XML (eXtensible Markup Language)
  - resource for more information: http://sbml.org

SBML software and models

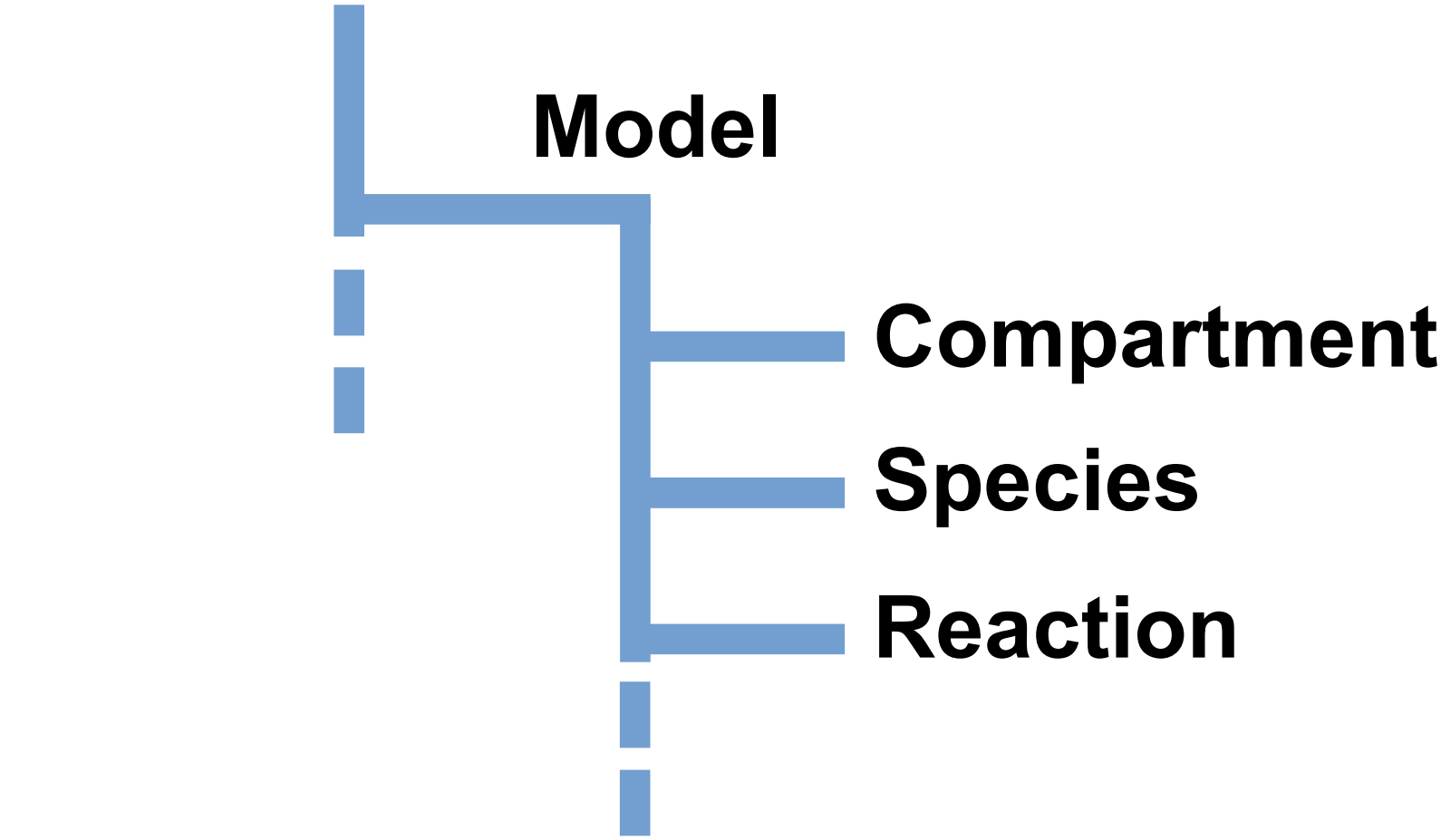
> 250 software tools support SBML



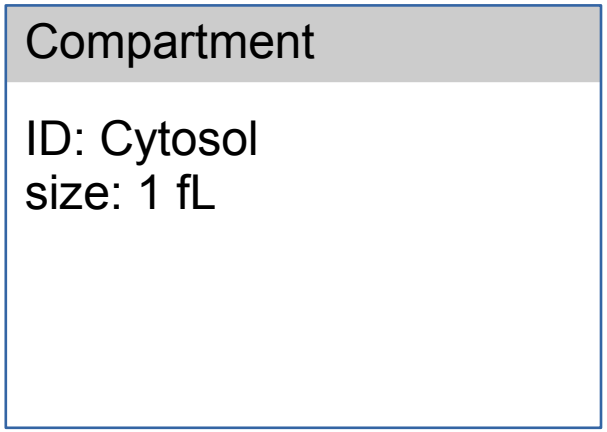
Raise of the amount of SBML tools over the last decade

> 1000 SBML models available

SBML Structure

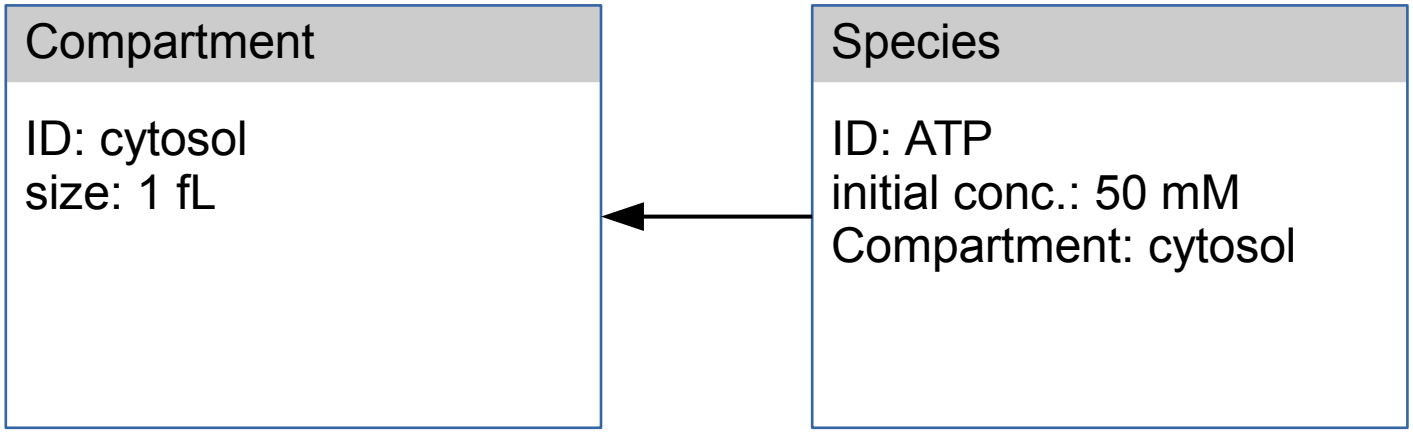


SBML Structure



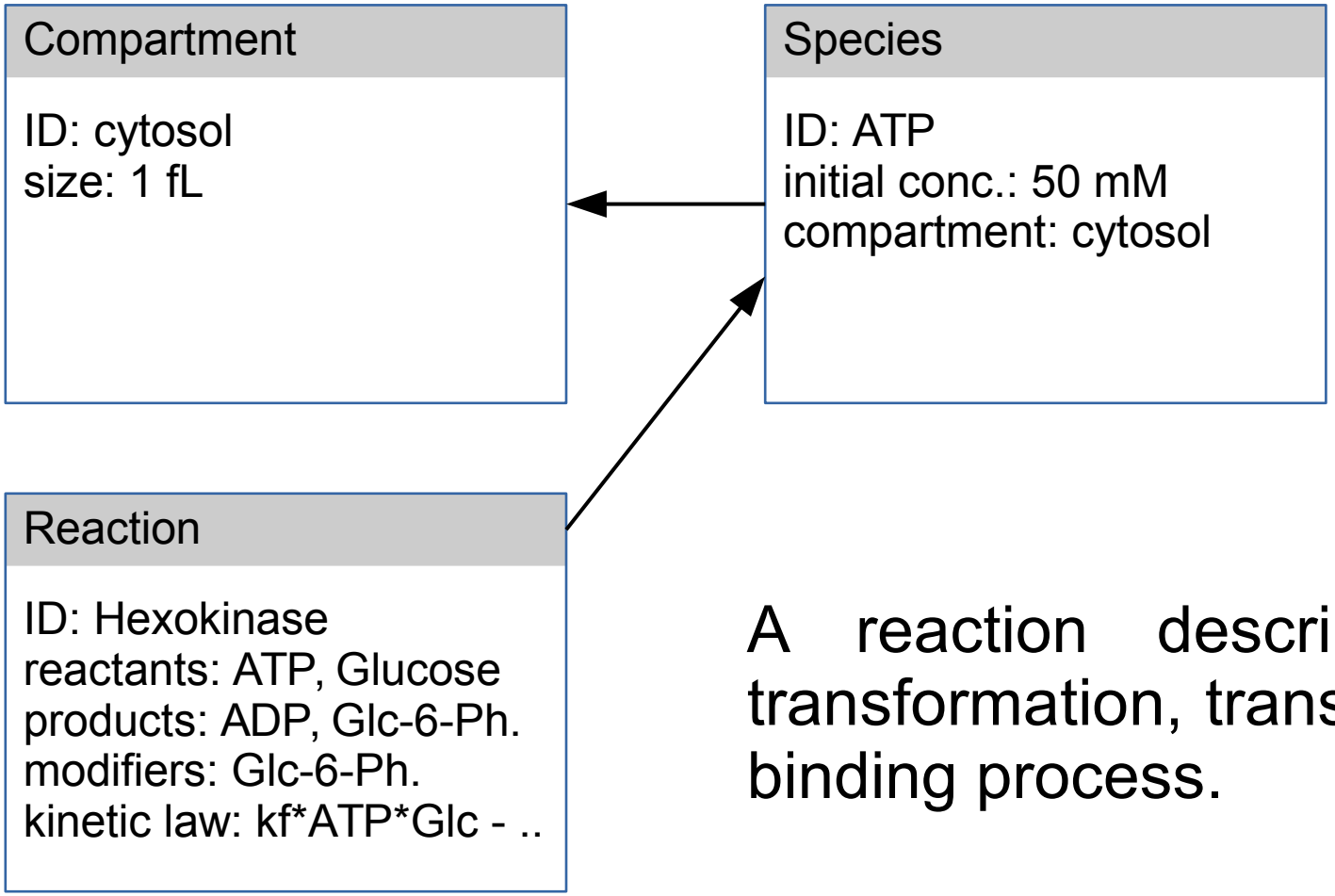
A compartment is a **container** that has **a defined size** (volume or area in case of membranes), for instance the cytosol or cell membrane

SBML Structure



A **species** defines a biochemical entity that can participate in reactions, for instance Glucose-6-phospate

SBML Structure



A reaction describes a transformation, transport or binding process.

XML

- eXtensible Markup Language
- standard for storing textual information
- meta format that can be user defined
- examples:
  - OpenDocument (OpenOffice)
  - XHTML
  - SVG

XML elements

- XML element with start end end tag
  (a tag starts with < and ends with >)

```
<message>
        Hello World
</message>
```

- empty XML element

```
<message/>
```

- XML attributes

```
<message sender=”Jannis” />
```

XML has a treee structure

- Nesting of XML elements

```
<bookshelf location="hallway" >
    <color>blue </color>
    <book title="Momo" author="Michael E" >
        <p>In old, old times, when the ...</p>
    </book>
    <book title="Pippi Longst" author="Astrid L" >
        <p>At the end of a little ...</p>
    </book>
</bookshelf>
```

XML elements vs attributes

- XML elements

```
<person>
      <gender>female</gender>
      <firtname>Anna</firstname>
      <lastname>Smith</lastname>
</person>
```

- XML attributes

```
<person gender="female">
      <firtname>Anna</firstname>
      <lastname>Smith</lastname>
</person>
```

Right and wrong

- Well-formed

```
<bookshelf location="hallway" >
        <color>blue </color>
</bookshelf>
```

- Not well-formed

```
<bookshelf location="hallway" >
         <color>blue
</bookshelf>
         </color>
```

SBML code

SBML compartment

```
<listOfCompartments>
    <compartment id="cytosol" size="1e-14"/>
</listOfCompartments>3
```

SBML species

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initialAmount="0" name="ES"/>
    <species compartment="cytosol" id="P" initialAmount="0"  name="P"/>
    <species compartment="cytosol" id="S" initialAmount="1e-20" name="S"/>
    <species compartment="cytosol" id="E" initialAmount="5e-21" name="E"/>
</listOfSpecies>
```

Full SBML model

```
<?xml version="1.0" encoding="UTF-8"?>
 <sbml xmlns="http://www.sbml.org/sbml/level2/version1" version="1" level="2" >
  <model id="model" >
   <listOfCompartments >
    <compartment size="1" id="GO 0005623" name="cell" >
   </listOfCompartments >
   <listOfSpecies >
    <species initialConcentration="1" compartment="GO 0005623" id="ATP" name="ATP" >
    <species initialConcentration="1" compartment="GO 0005623" id="NADH" name="NADH" >
    <species initialConcentration="2" compartment="GO 0005623" id="ADP" name="ADP" >
    <species initialConcentration="2" compartment="GO 0005623" id="NADPH" name="NADPH" >
    <species initialConcentration="1" compartment="GO 0005623" id="enzyme" name="2.7.1.86" >
   </listOfSpecies >
   <listOfReactions >
    <reaction id="R00105" name="ATP NADH 2 phosphotransferase" >
     <listOfReactants >
      <speciesReference species="ATP" >
      <speciesReference species="NADH" >
     </listOfReactants >
     <listOfProducts >
      <speciesReference species="ADP" >
      <speciesReference species="NADPH" >
     </listOfProducts >
     <listOfModiers >
      <modierSpeciesReference species="enzyme" >
     </listOfModiers >
    </reaction >
   </listOfReactions >
  </model >
 </sbml>
```

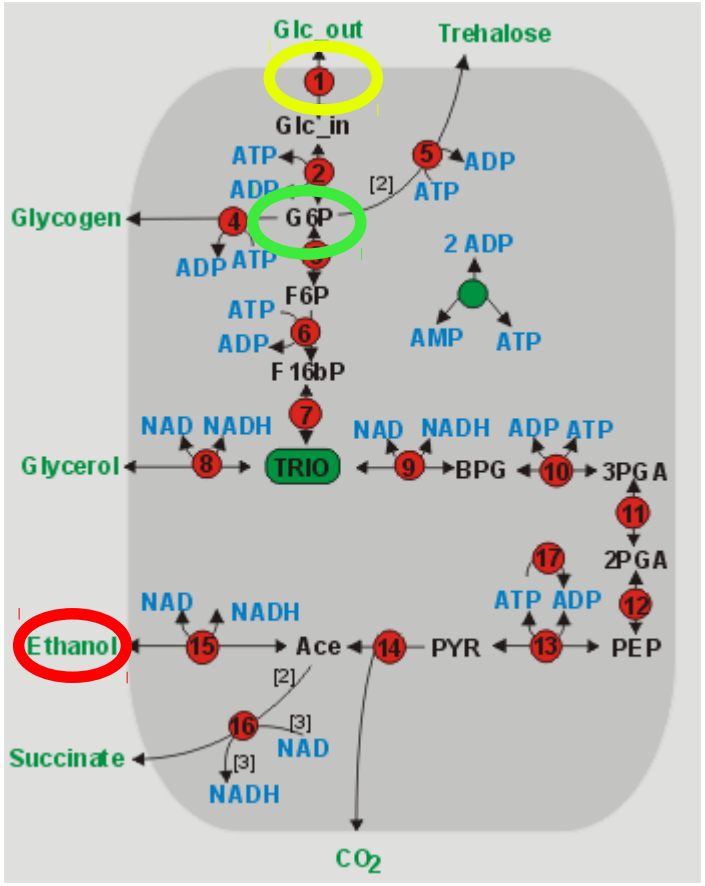
ATP + NADH ↔ ADP + NADPH

Model annotation

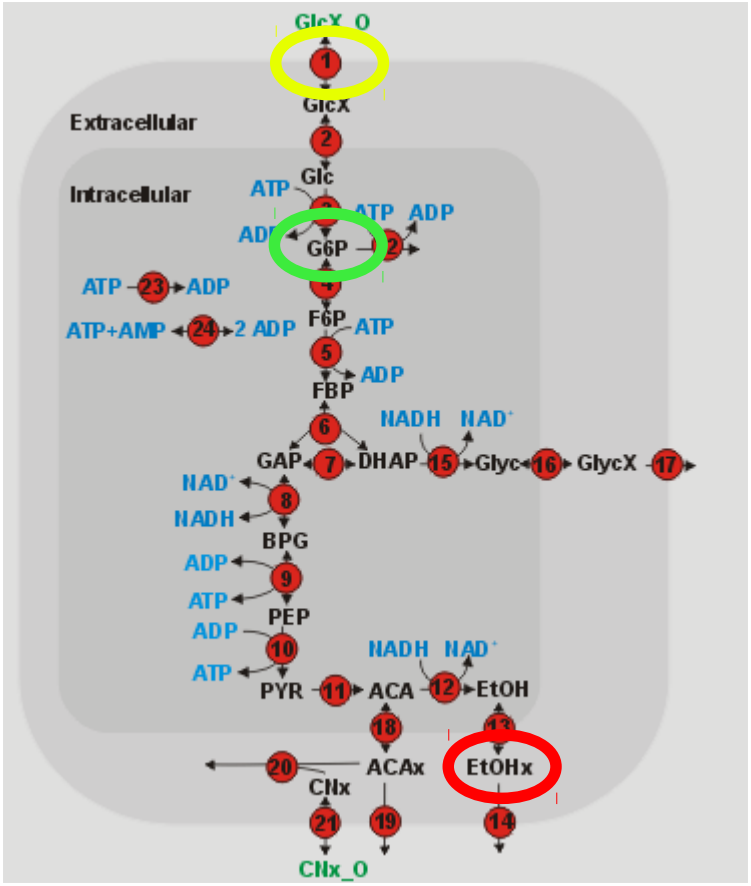
- Problem:
  - how to define the biological meaning of SBML elements?
  - e.g. **Glc** stands for Glucose?, D-Glucose? L-Glucose?

Model annotation

Hynne



Teusink



Model annotation

- Solution: MIRIAM annotations
  - link SBML elements to unique database entries

Subject	Predicate	Object
ATP	isVersionOf	Reactome ID 154722

- (some) Databases relevant for annotation:
  - KEGG (reactions, species)
  - GeneOntology (compartments, species, reactions)
  - ChEBI (small molecules)
  - UniProt (proteins)

MIRIAM qualifiers

Biology Qualifiers

- is
- hasPart
- isPartOf
- isVersionOf
- hasVersion
- isHomologTo
- isDescribedBy
- isEncodedBy
- encodes
- occursIn

Model Qualifiers

- is
- isDescribedBy

Graphic visualization - SBGN

SBGN is an effort to standardize

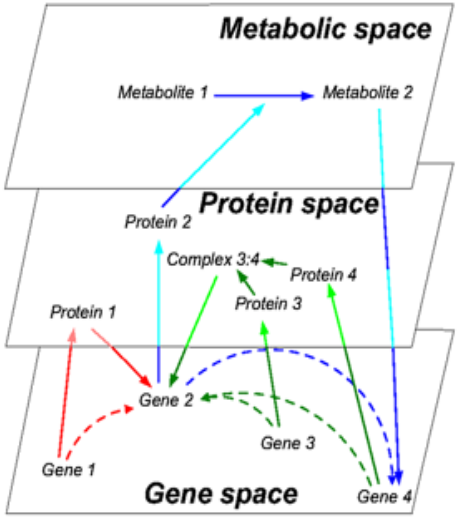
the graphical notation used in maps of

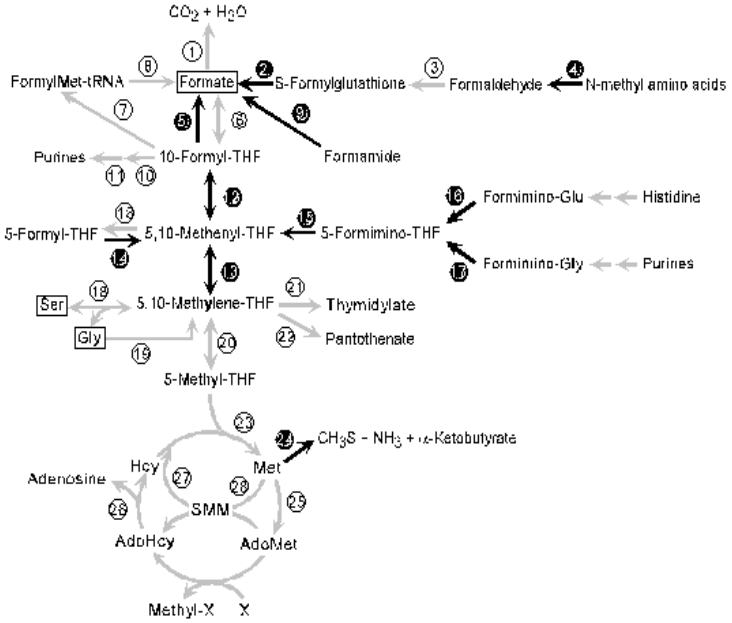
biochemical and cellular processes

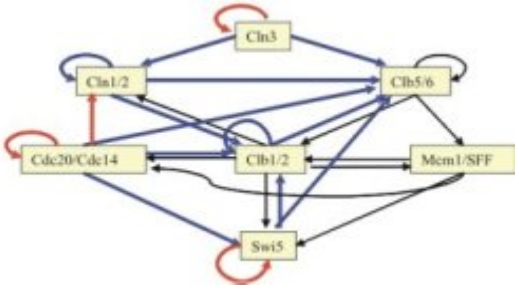
studied in Systems Biology.

www.sbgn.org

SBGN – Why is it needed?

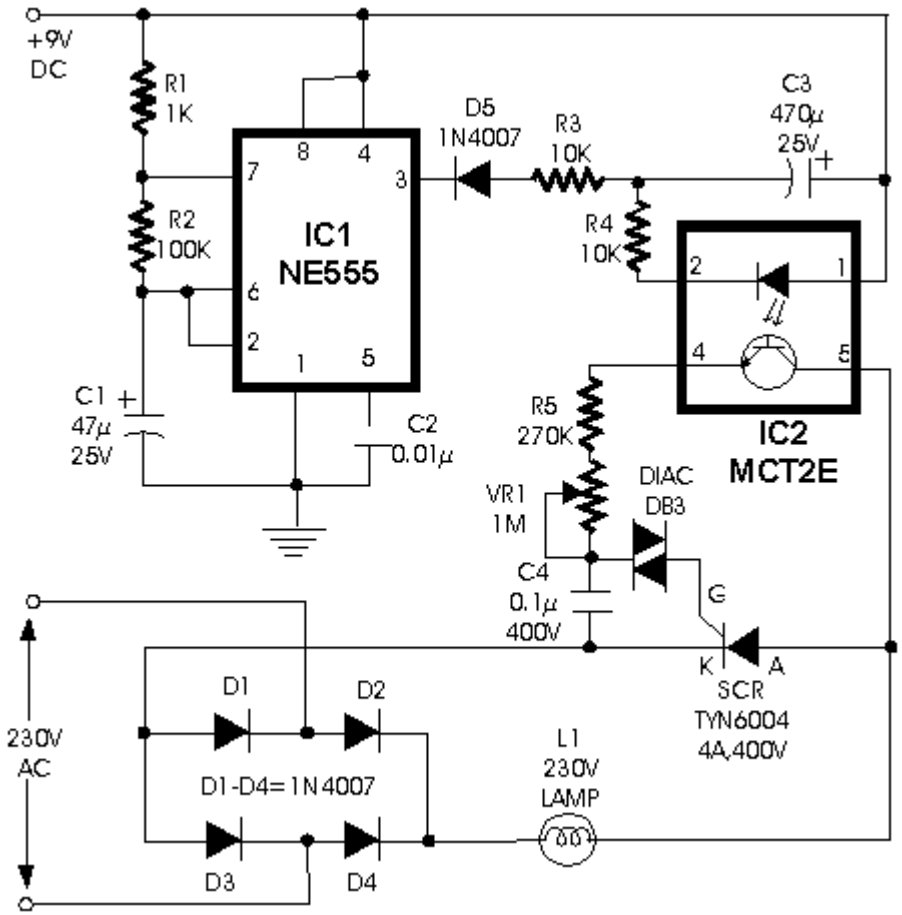




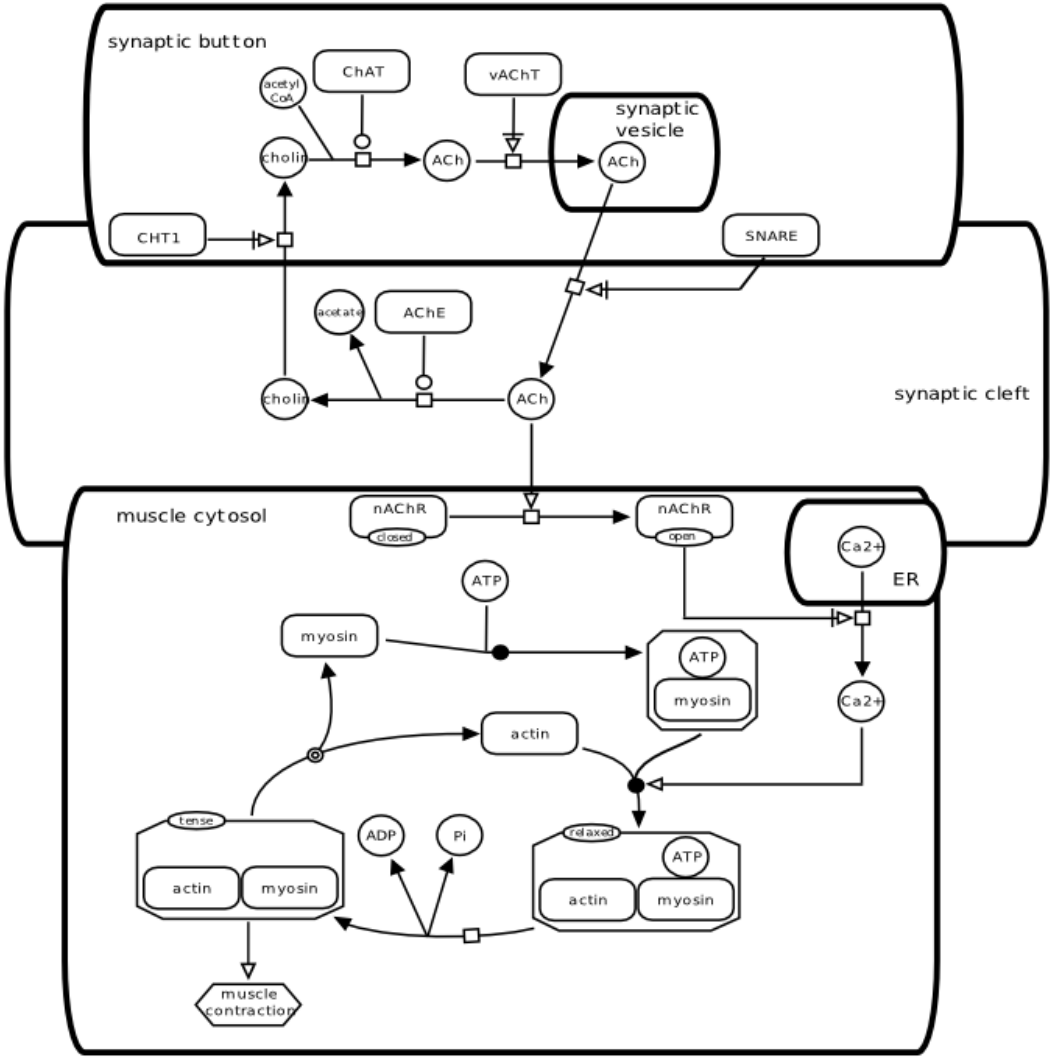


Too many different kinds of graphical notations for biochemical network diagrams

Electronic engineering has a universal graphical language



SBGN - The universal graphical language for Systems Biology

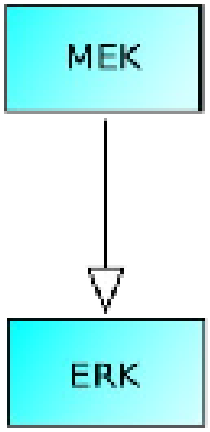


SBGN consists of 3 different languages

- Process Description Language
- Entity Relationship Language
- Activity Flow Language

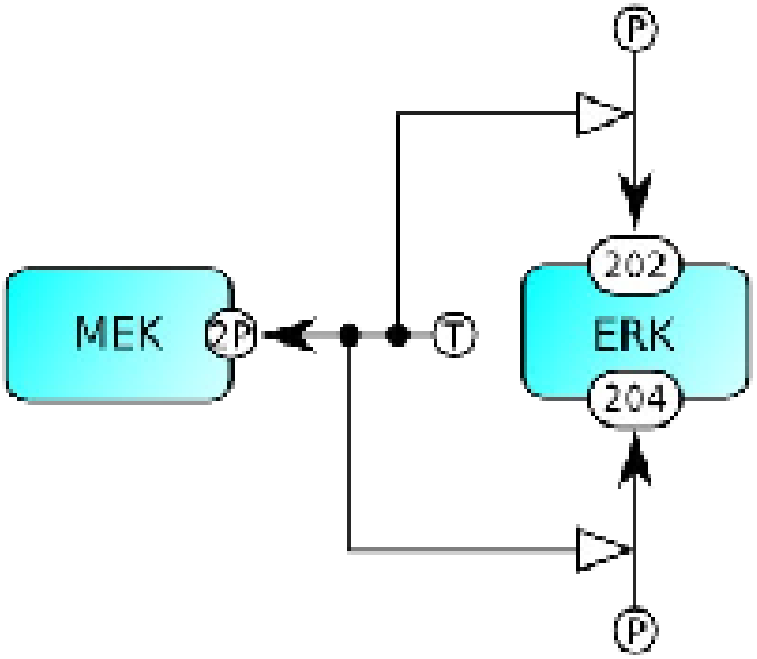
3 SBGN languages

1. Activity Flow Language (AF): depicts the information flow between entities in a network. Used to represent perturbation effects and describe coarse-grained interaction networks.



3 SBGN languages

2. Entity Relationship Language (ER): shows all relationships in which an entity participates, regardless of temporal aspects. Can be used to describe the molecular biology.



3 SBGN languages

3. Process Description Language (PD): shows temporal courses of biochemical interactions. An entity can appear multiple times to visualize all molecular interactions.