



# Annot package

Neil Swainston  
Manchester Centre for Integrative Systems Biology

3 May 2010



Annot package

“SBML should not pre-emptively incorporate features that could be useful `at some point’”



## Annot package

- Annot package workshop later this month...
  - [http://sbml.org/Events/Other\\_Events/Annotation\\_package\\_workshop\\_2010](http://sbml.org/Events/Other_Events/Annotation_package_workshop_2010)
- Questions for today:
  - **why** we do / don't annotate models?
  - **what** we would like to do that we currently cannot, either due to lack of tools or more seriously, lack of support in SBML?
  - things that we would like to do that rely on annotations, but we **don't realise it yet?**
- Less fussed by “how”

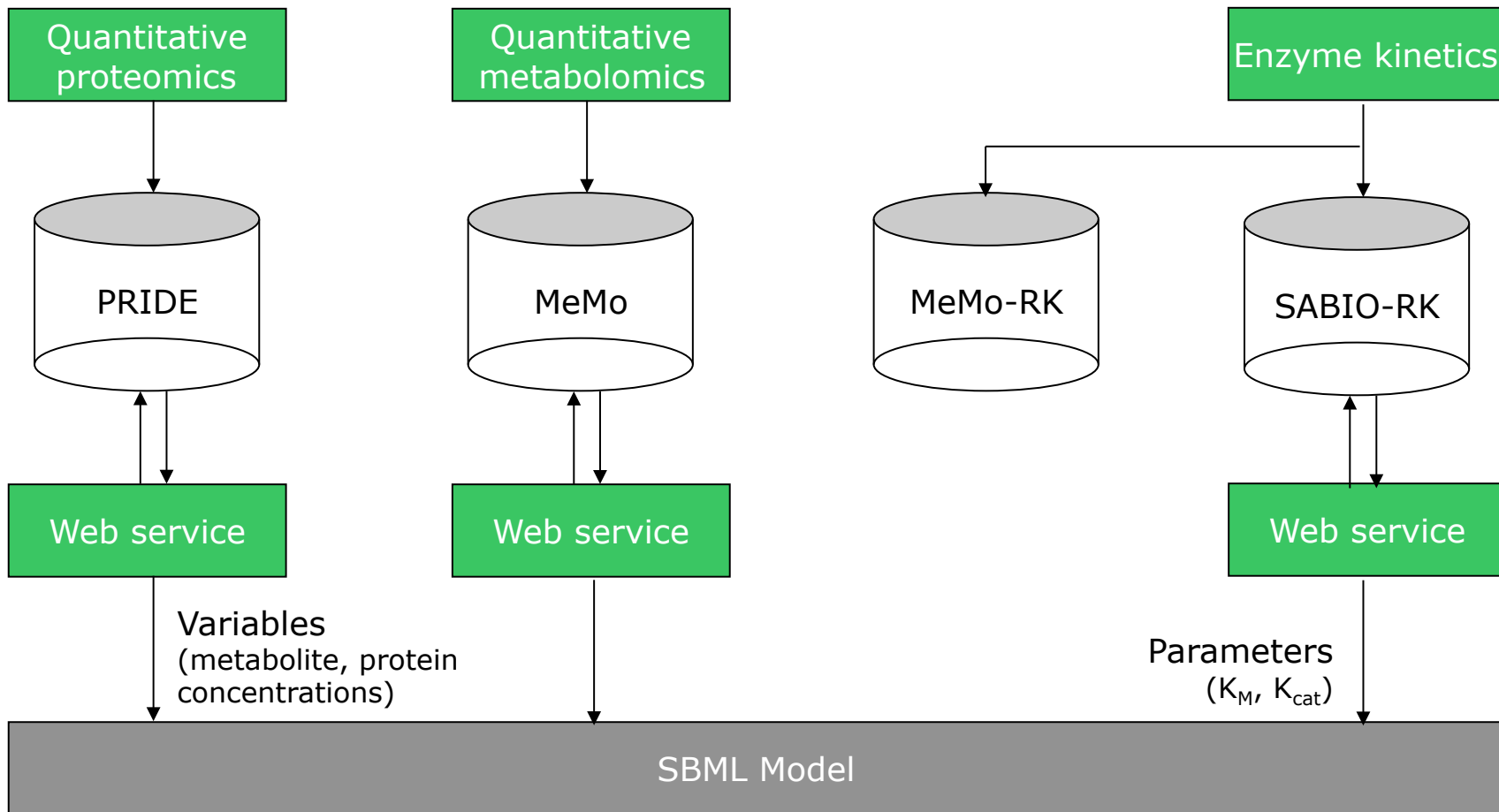


## Why I annotate

- Provide **definitive, unique identifier** for “stuff”
  - One motivation: mapping experimental data to models
  - Need **annotated models AND annotated data**



# Experimental project overview





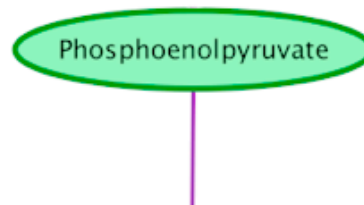
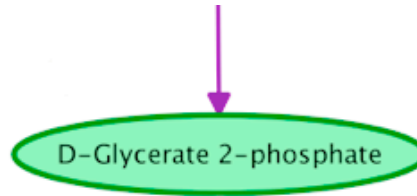
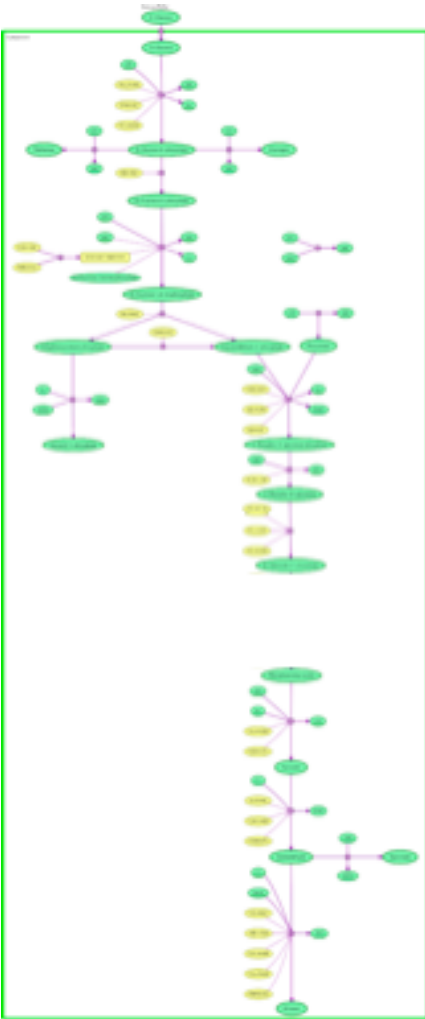
## Genome-scale models

- **Ridiculous-ome** – Jim Schaff
- Genome-scale SBML model of metabolism
  - Requires annotation to understand / query (bloody huge)
  - Requires annotation to build (merged from smaller models)



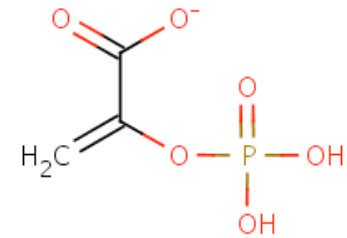
## Purpose of annotations

- More interesting purpose...
  - Use the annotations to **reason** over the model
  - Does my model accurately reflect the biochemistry?
  - Do my reactions **elementally balance**? Are the **stoichiometries** correct?
  - Easy to determine if we “**deference**” ChEBI identifier (or whatever) to determine chemical formulae

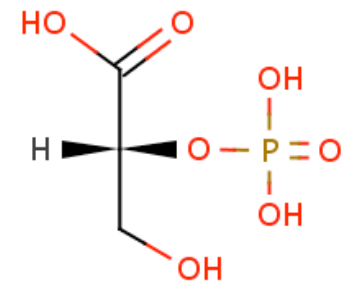


## Gap-filling

CHEBI:17835



CHEBI:18021





## Limitations

- We can only annotate **elements** not **attributes**...

A screenshot of an XML editor window titled 'BIOMD0000000003.xml'. The editor shows the following XML code:

```
1 <species metaid="_metal" id="id1" name="Lipopolysaccharide"
2   initialConcentration="50.2" units="nM"
3   compartment="EndoplasmicReticulum">
4   <annotation>
5     <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
6       xmlns:bqbiol="http://biomodels.net/biology-qualifiers/">
7       <rdf:Description rdf:about="#_metal">
8         <bqbiol:isDescribedBy>
9           <rdf:Bag>
10            <rdf:li rdf:resource="urn:miriam:obo.eco:ECO%3A0000123"/>
11            <rdf:li rdf:resource="urn:miriam:pubmed:17882155"/>
12          </rdf:Bag>
13        </bqbiol:isDescribedBy>
14      </rdf:Description>
15    </rdf:RDF>
16  </annotation>
17 </species>
```



## Limitations

- Who did what when?
  - Improved history
- Annotations of annotations
  - X was annotated with Y because of publication Z
- Examples being collected here...
- [http://docs.google.com/View?id=dftdk3rc\\_58gvv857tg](http://docs.google.com/View?id=dftdk3rc_58gvv857tg)
- Speak to me...



## Versioning?

- Species X in Model Y, version 1 was Species A in Model B, version 6
- Is this an annotation issue?



## Annot package

- Questions for today:
  - **why** we do / don't annotate models?
  - **what** we would like to do that we currently cannot, either due to lack of tools or more seriously, lack of support in SBML?
  - things that we would like to do that rely on annotations, but we **don't realise** it **yet**?



# Annot package

Neil Swainston  
Manchester Centre for Integrative Systems Biology

3 May 2010