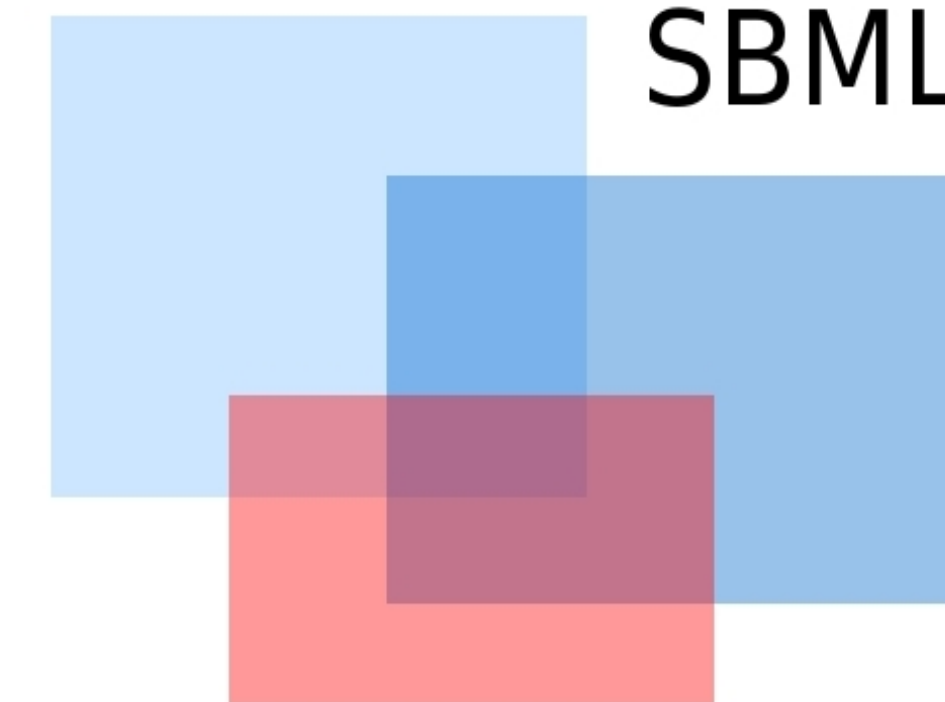


SemanticSBML: computer-assisted construction, checking, and merging of biochemical models

Wolfram Liebermeister and Edda Klipp
{ lieberme | klipp } @molgen.mpg.de

Semantic
SBML



Computational Systems Biology, Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany

Overview

SBML¹ is a popular XML-based standard format for mathematical models in systems biology. As mathematical models in cell biology are growing in number and size, flexible ways to combine models will gain additional importance; model combination requires that the meaning of model elements is specified by annotations.

The SemanticSBML tools assist the user in annotating, checking, and combining SBML models: SBMLannotate helps the user to annotate SBML model elements with unique identifiers pointing to databases such as KEGG or Gene Ontology. SBMLcheck verifies the syntactic and semantic validity of SBML files. SBMLmerge helps the user to merge SBML models. In the future, we envisage to support annotations for protein complexes, phosphorylation states, substance classes, and lumped reactions.

The SemanticSBML tools can be freely downloaded and accessed via a web interface at <http://sysbio.molgen.mpg.de/semanticsbml>.

SBMLannotate

SBMLannotate allows you to link the elements of an SBML model to unique identifiers (IDs) from different databases (currently GO, KEGG, ChEBI, PubChem, 3DMET, and CAS; we envisage to support Reactome, Uniprot, MetaCyc). Users can also include new databases into SBMLannotate.

The annotations are stored in the SBML code in a MIRIAM³-compliant data format⁴.

```
<species metaid="atp" id="atp" compartment="cell" initialConcentration="1.31">
  <annotation>
    <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dc="http://purl.org/dc/elements/1.1/">
      <dc:relation>
        <rdf:Bag>
          <rdf:li rdf:resource="http://www.cas.org/#56-65-5"/>
          <rdf:li rdf:resource="http://pubchem.ncbi.nlm.nih.gov/#3384"/>
          <rdf:li rdf:resource="http://www.genome.jp/kegg/compound/#C00882"/>
          <rdf:li rdf:resource="http://www.ebi.ac.uk/chebi/#15422"/>
        </rdf:Bag>
      </dc:relation>
    </rdf:RDF>
  </annotation>
</species>
```

SBML species tag "ATP" with annotations

For automatic annotation, SBMLannotate screens the SBML code for known names and IDs and suggests appropriate annotations; the user can also retrieve names and database IDs via a string search. To annotate a chemical reaction, SBMLannotate searches the KEGG² database for reactions containing the respective metabolites as substrate and products. It is also possible to create entire annotated SBML models from a list of KEGG reaction IDs.

Biochemical models in SBML format



SBMLannotate

helps the user to annotate model elements (compartments, metabolites, and reactions) with unique identifiers from various databases.

SBMLcheck

checks whether the SBML model contains correct and consistent annotations.

SBMLmerge

merges SBML models into a single consistent model.

Merged SBML file

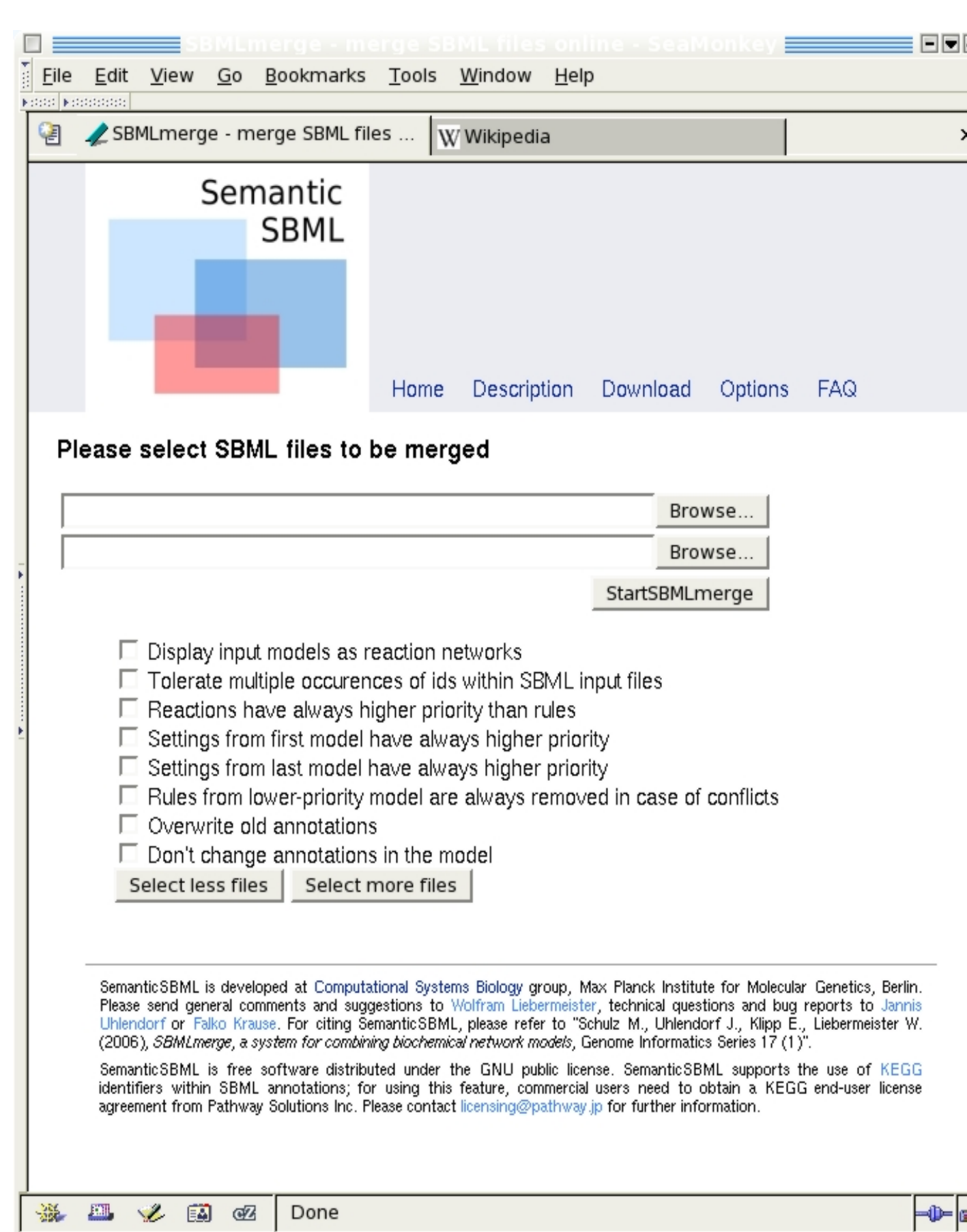


SBMLcheck

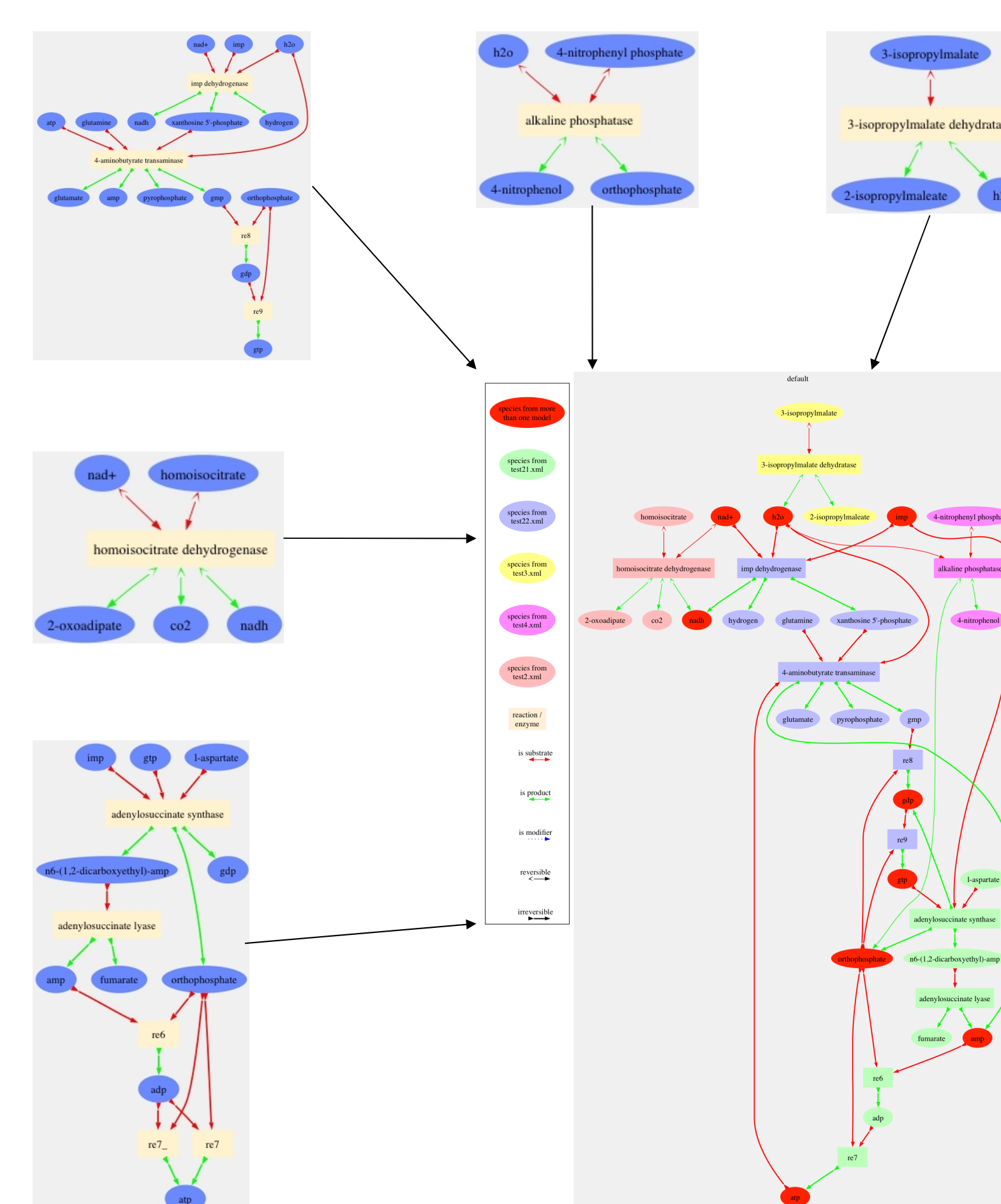
SBMLcheck performs various checks on the input SBML files. First of all, it verifies whether the model contains readable and consistent annotations. Elements without annotations are reported to the user. Furthermore SBMLcheck detects mistakes in the model structure, such as circular dependencies among algebraic equations or violations of atom number conservation in chemical reactions.

SBMLmerge

SBMLmerge⁶ assists you in merging two or more annotated models: during merging, it detects and resolves various syntactic and semantic conflicts, such as conflicting variable names, elements that appear in both input models, and also mathematical problems arising from the combination of equations. First, the program screens the models for identical elements. If the input models make contradicting statements about the same biochemical quantity, the user is asked to choose one of them. Afterwards, the program resolves circular dependencies in assignment rules, which may arise by the user's choices. In the end, the merging process results in a new, valid SBML model that can be visualised as a reaction network graph⁵.



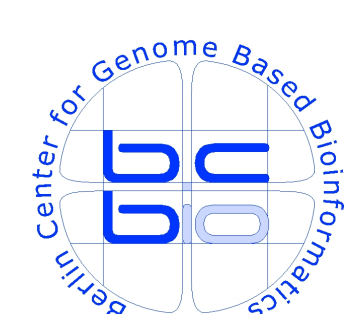
Web interface of SBMLmerge



Models merged by SBMLmerge

Financial support:

This work was financed by Berlin Center for Genome based Bioinformatics, BCB (German Federal Ministry for Education and Research, grant 031U109C and 0312705B); European Commission, grant No. 503269 (EMI-CD); and the European integrated project BaSysBio.



References

- [1] Finney *et al.* (2003) **Systems Biology Markup Language (SBML) Level 2: Structures and Facilities for Model Definitions**
- [2] Kanehisa *et al.* (2006) **From Genomics To Chemical Genomics: New Developments KEGG (Nucleic Acid Res.) 34, D354 - 357**
- [3] LeNovere *et al.* (2005) **Minimum information requested in the annotation of biochemical models (MIRIAM) (Nat. Biotechnol.)**
- [4] LeNovere and Finney (2005) **A simple scheme for annotating SBML with references to controlled vocabularies and database entries**
- [5] <http://www.graphviz.org>
- [6] Schulz *et al.* (2006), **SBMLmerge, a system for combining biochemical network models**, *Genome Informatics Series 17 (1)*