Research Coordination Network for Systems and Synthetic Biology Standards

Chris J. Myers

University of Utah

August 22, 2014

Research Coordination Network

- In June, we submitted a proposal to NSF to establish a *research coordination network* (RCN) for systems and synthetic biology standards.
- The goal of this RCN is to facilitate the synthetic biology community in joining the system biology's COMBINE organization.
- RCN steering committee:
  - Chris Myers (Utah), Chair
  - Gary Bader (Toronto)
  - Douglas Densmore (Boston)
  - Mike Hucka (Caltech)
  - Nicolas Le Novère (Babraham)
  - Jacqueline Quinn (Google)
  - Herbert Sauro (Washington)
  - Falk Schreiber (Monash)
  - Dagmar Waltemath (Rostock)
  - Anil Wipat (Newcastle)

Focus of Research Coordination

- Specification infrastructure
- Metadata and annotations
- COMBINE archive
- SBOL Visual/SBGN
- SBOL/SBML/SED-ML
- SBOL/BioPAX
- Repositories
- Journals
- Compliance testing

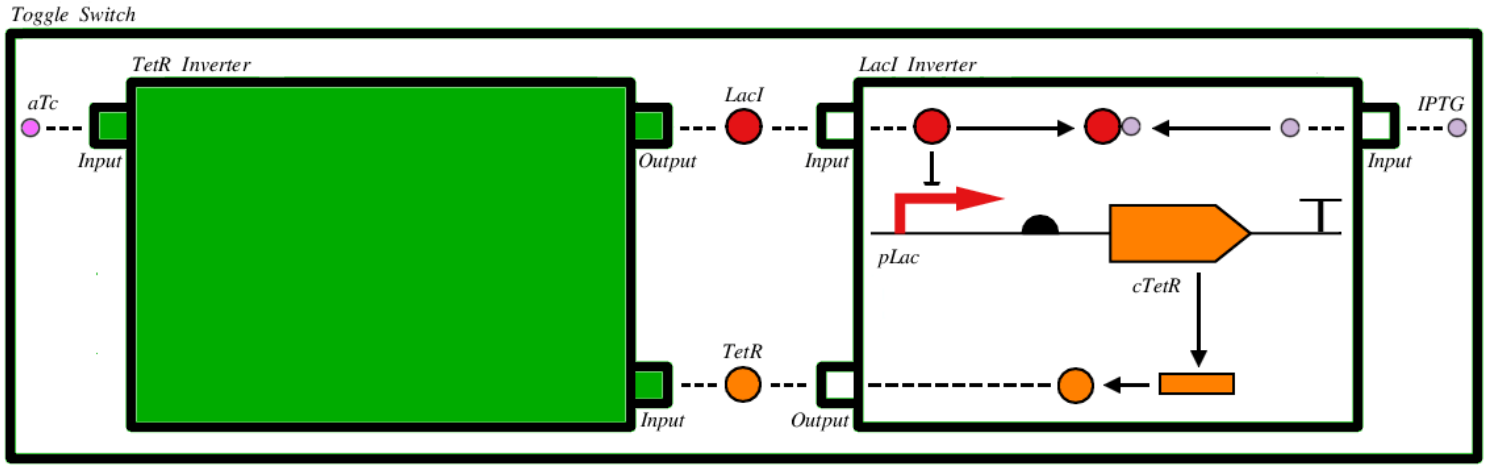
Focus of Research Coordination

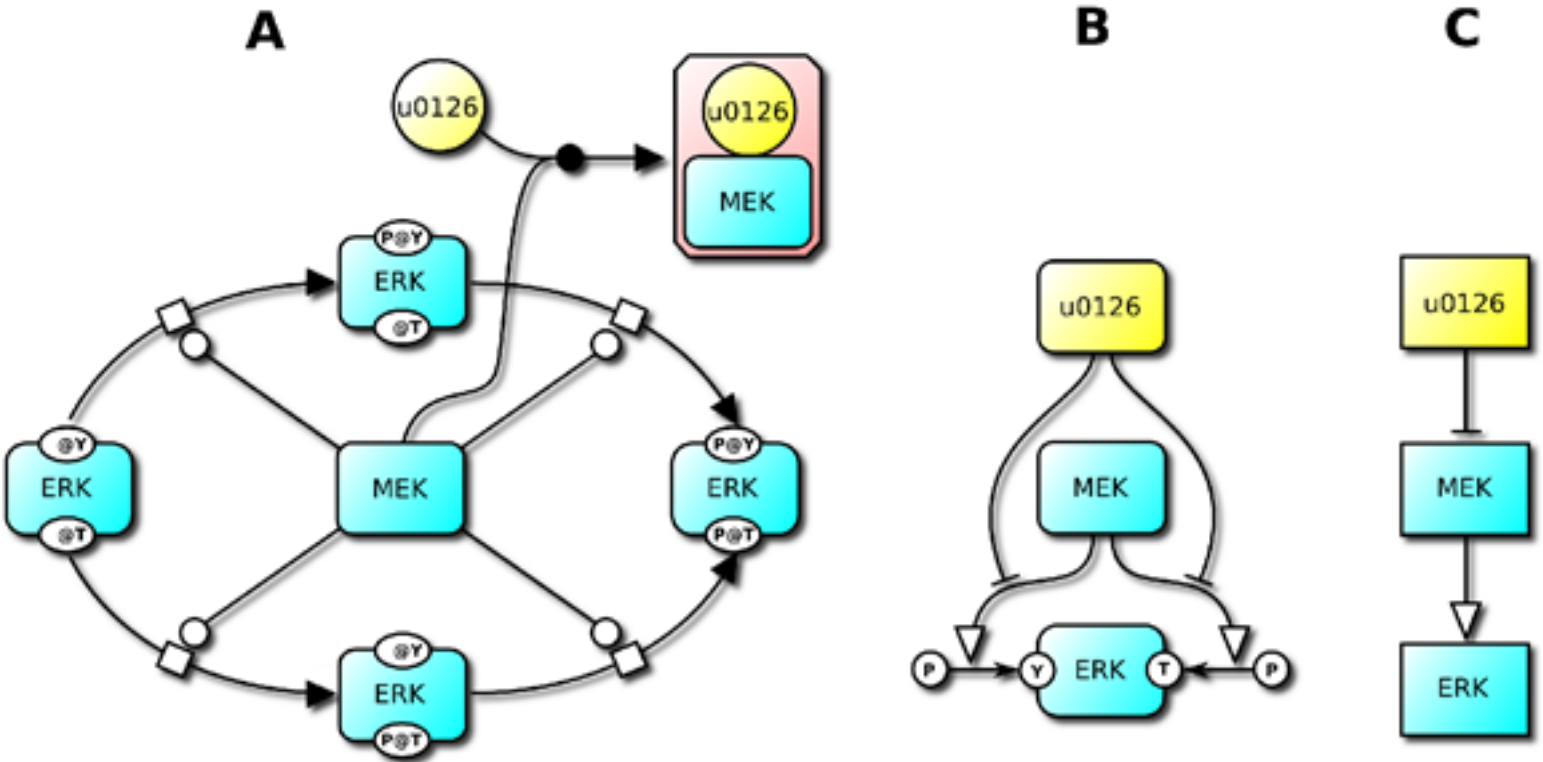
- Specification infrastructure
- Metadata and annotations
- COMBINE archive
- SBOL Visual/SBGN
- SBOL/SBML/SED-ML
- SBOL/BioPAX
- Repositories
- Journals
- Compliance testing

Focus of Research Coordination

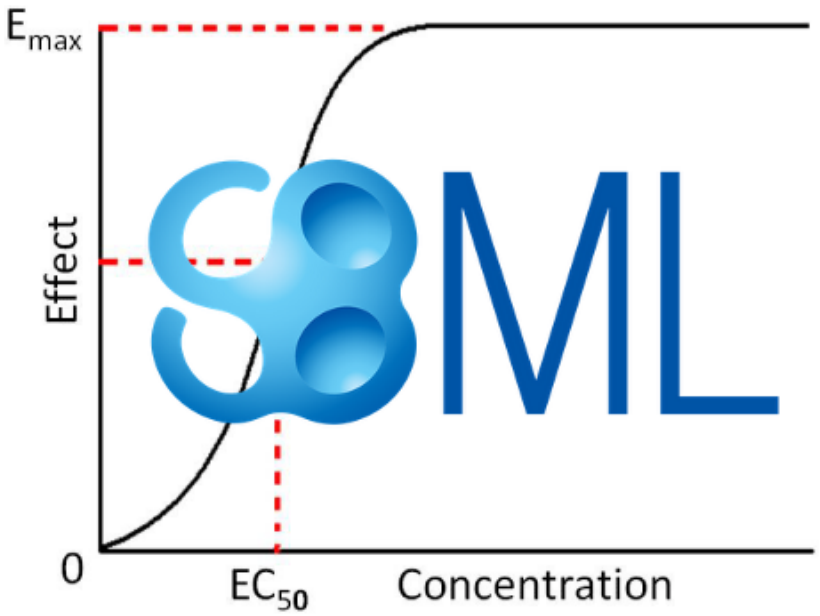
- Specification infrastructure
- Metadata and annotations
- COMBINE archive
- SBOL Visual/SBGN
- SBOL/SBML/SED-ML
- SBOL/BioPAX
- Repositories
- Journals
- Compliance testing

SBOL Visual/SBGN



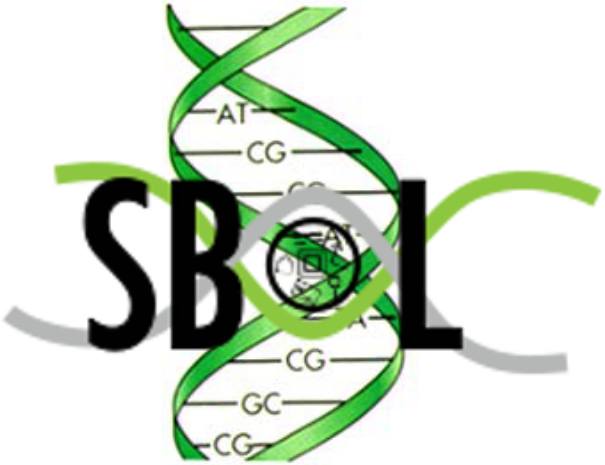


Connecting SBML to SBOL



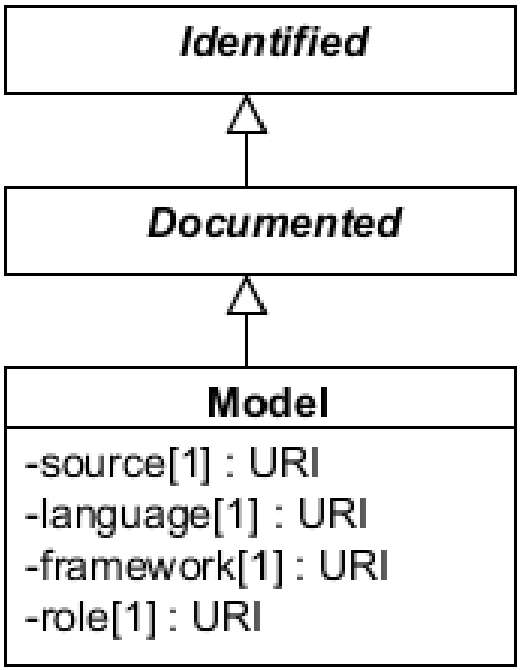
SBML

RDF/XML
Annotation



Roehner et al., ACS Synthetic Biology (2013).

Connecting SBOL to SBML

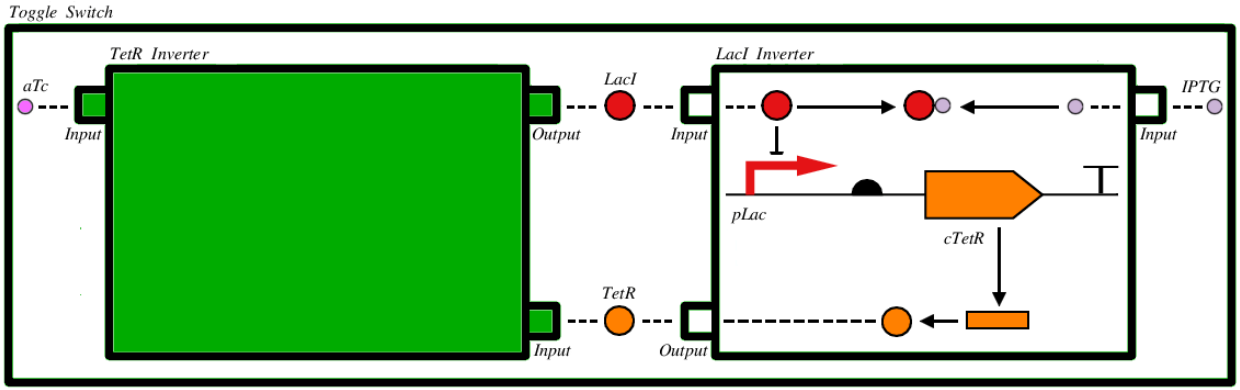


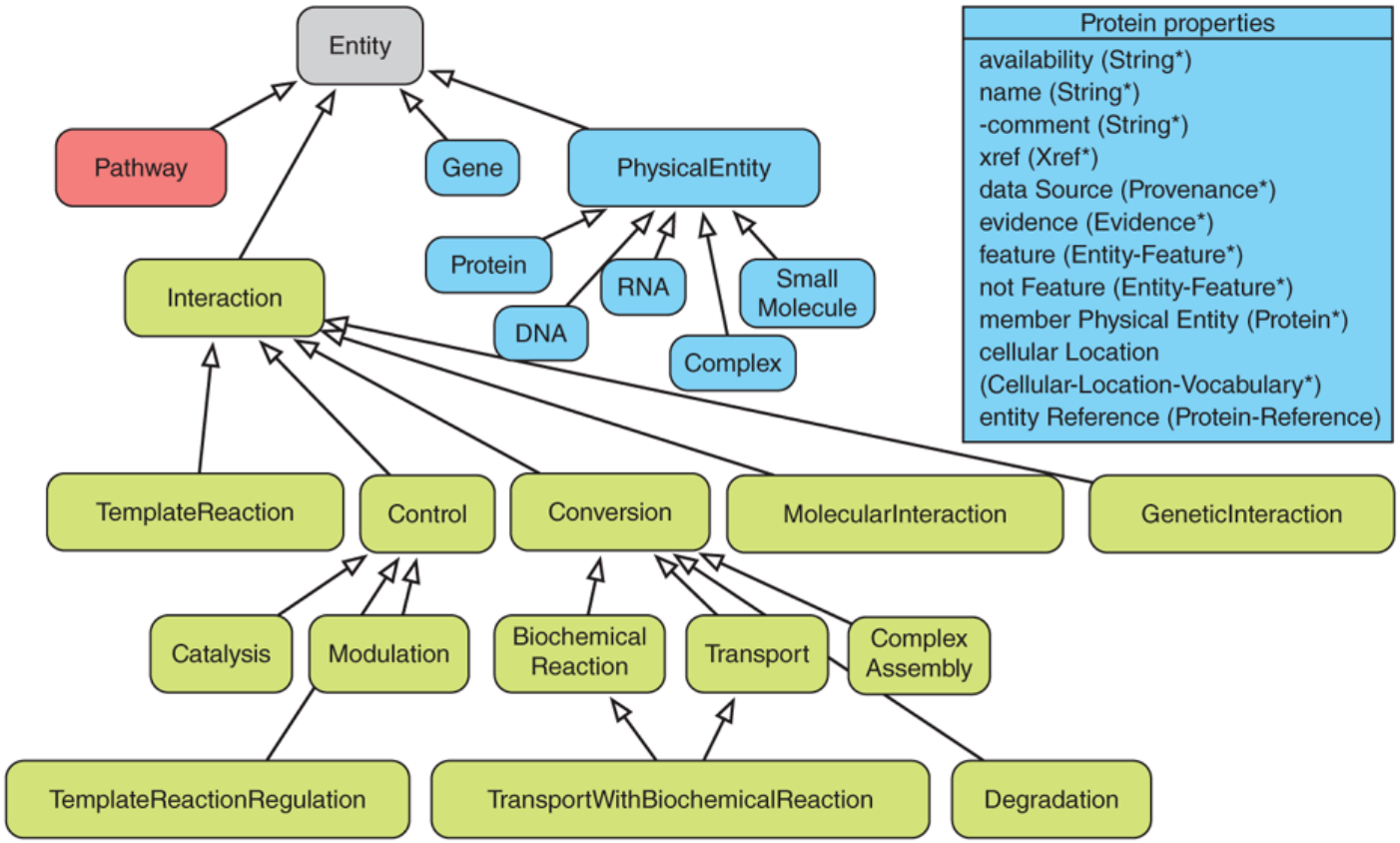
- *Source* is pointer to model file location.
- *Language* may be SBML, CellML, Matlab, etc.
- *Framework* is modeling formalism (ODE, stochastic, Boolean, etc.).
- *Role* is the purpose of the model.

SBOL/SBML/SED-ML: Discussion

- Are we happy with the methods proposed for connecting SBOL and SBML models?
- Where do we draw the line between SBOL functional information and SBML models?
- Should we be translating SBOL's functional information into SBML models or vice versa?

SBOL/BioPAX


Toggle Switch
TetR Inverter
LacI Inverter
aTc
Input
Output
LacI
IPTG
pLac
cTetR
TetR


Entity
Pathway
Gene
PhysicalEntity
Protein
RNA
Small Molecule
DNA
Complex
Interaction
TemplateReaction
Control
Conversion
MolecularInteraction
GeneticInteraction
Catalysis
Modulation
Biochemical Reaction
Transport
Complex Assembly
TemplateReactionRegulation
TransportWithBiochemicalReaction
Degradation
Protein properties
availability (String*)
name (String*)
-comment (String*)
xref (Xref*)
data Source (Provenance*)
evidence (Evidence*)
feature (Entity-Feature*)
not Feature (Entity-Feature*)
member Physical Entity (Protein*)
cellular Location
(Cellular-Location-Vocabulary*)
entity Reference (Protein-Reference)

SBOL/BioPAX: Discussion

- Can translators be developed to extract interactions from BioPAX?
- Can interactions in SBOL be converted into BioPAX?
- Should SBOL refer to BioPAX records for the information eliminating the need for these interactions to be stored within the SBOL file at all?

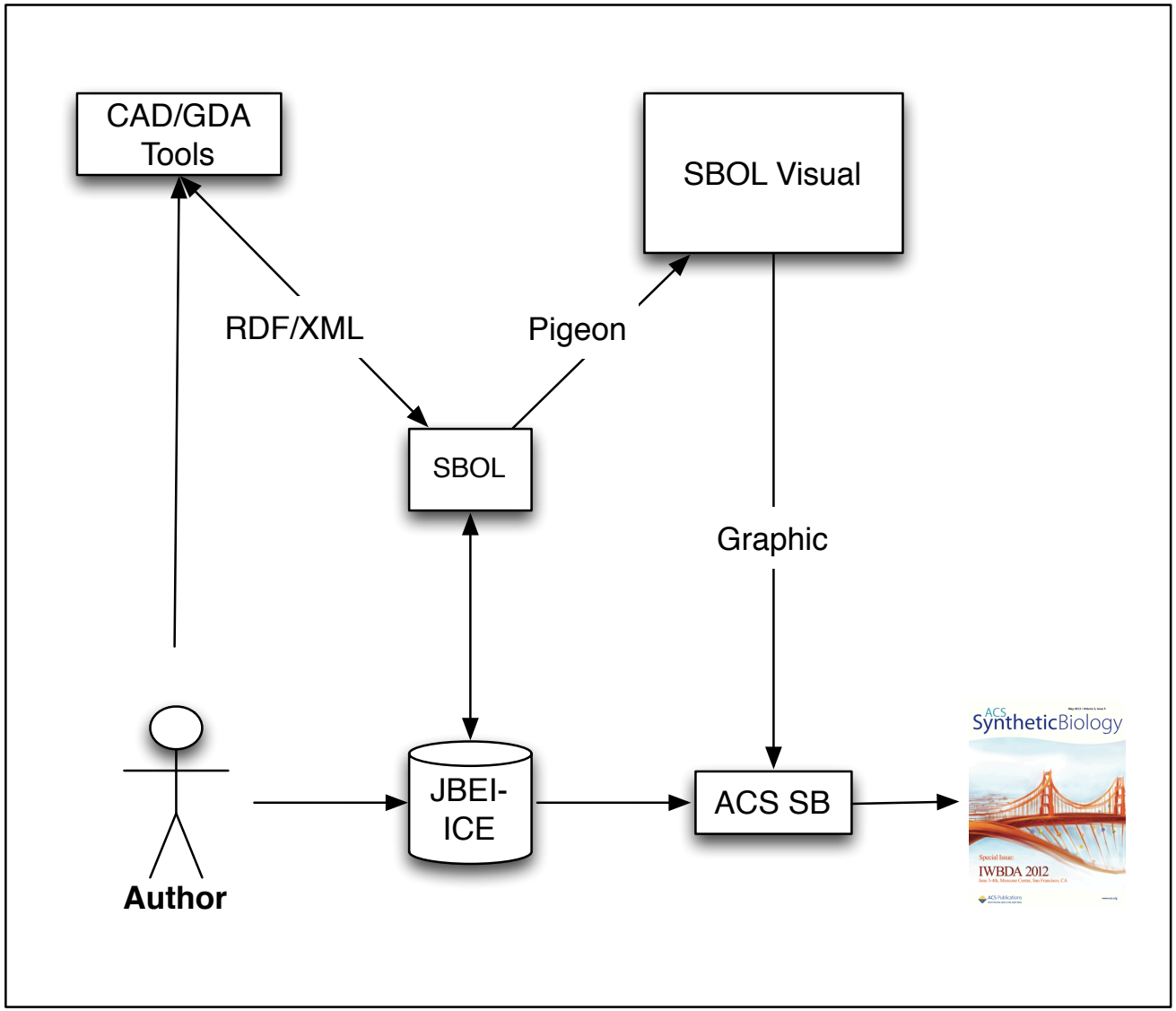
Repositories: Discussion

- Many repositories for separate standards:
  - Pathway Commons (BioPAX)
  - BioModels (SBML)
  - iGEM Registry, JBEI-ICE, SBOL-Stack (SBOL)
  - Virtual Parts (SBOL/SBML)
- It would be desirable to have a single interface enabling a user to obtain all the necessary data from one information portal.
- How might this work?

Journals

- The use of standards for DNA sequence data became commonplace when journals began to require them for publication.
- While there are journals that encourage the use of standards, such as SBML, for modeling, there are currently no journals that require it.
- In order to encourage journals to require these standard data representations, the impact on the authors must be minimized.
- We should have a user friendly portal to their repositories that enables authors to easily deposit their models and designs.
- These interfaces should allow authors to provide their information in an intuitive way while storing their information using an appropriate standard.
- This task is perhaps the one that has the potential to have the largest overall impact on the community.

SBOL and ACS Synthetic Biology



Journals: Discussion

- How do we convince journals to encourage (and ideally require) the use of standards?
- How can we make it easier for authors to use standards in their journal publications?
- What infrastructure is needed to support this effort?

Compliance Testing

- The challenge:
  - A standard should be broad enough to support users needs.
  - A standard should not be so broad that no tools support the standard.
- If a standard is too light, users complain they cannot encode their data.
- If a standard is too heavy, users complain that tools do not interoperate.
- In both cases, the standard is blamed, but this is really a tool problem.
- A possible solution is *compliance testing*.

Compliance Testing: Discussion

- What does it mean for a tool to support a standard?
- How do we encourage tool developers to fully support a standard?
- How do we collect data on the degree of standard support by tools?
- How do we advertise this information?
- Should COMBINE sanction workflows known to work?

Final Question

- Are there other areas where systems and synthetic biology can coordinate their efforts?