

The Synthetic Biology Open Language (SBOL)

Chris J. Myers
on behalf of the SBOL communities

Synthetic Biology

- *Synthetic biology* extends *genetic engineering* by adding:
 - *Standards*
 - *Abstraction*
 - *Decoupling*

(source: Drew Endy)

- *Synthetic biology* extends *genetic engineering* by adding:
 - *Standards*
 - *Abstraction*
 - *Decoupling*

(source: Drew Endy)

Essential information for synthetic DNA sequences

To the Editor:

Following a discussion by the workgroup for Data Standards in Synthetic Biology, which met in June 2010 during the Second Workshop on Bidesign Automation in Anaheim, California, we wish to highlight a problem relating to the reproducibility of the synthetic biology literature. In particular, we have noted the very small number of articles reporting synthetic gene networks that disclose the complete sequence of all the constructs they describe.

To our knowledge, there are only a few examples where full sequences have been released. In 2005, a patent application¹ disclosed the sequences of the toggle switches published four years earlier in a paper by Gardner *et al.*². The same year, Basu *et al.*³ deposited their construct sequences for programmed pattern formation into GenBank³. Examples of synthetic DNA sequences derived from standardized parts that have been made available in GenBank include the refactored genome of the bacteriophage

gaps between key components are almost never reported, presumably because they are not considered crucial to the report. Yet, synthetic biology relies on the premise that synthetic DNA can be engineered with base-level precision.

Missing sequence information in papers hurts reproducibility, limits reuse of past work and incorrectly assumes that we know fully which sequence segments are important. For example, many synthetic biologists are currently realizing that translation initiation rates are dependent on more than the Shine-Dalgarno sequence⁴. Sequences upstream of the

start codon are crucial for translation rates, yet are underreported. Similarly, it has been demonstrated that intron length can affect the dynamics of genetic oscillators⁵. Many more such examples are likely to emerge.

Because full sequence disclosure is critical, we wonder why the common requirement by many journals to provide GenBank entries

for genomes and natural sequences has

and welcome contributions from the greater community.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Jean Peccoud¹, J Christopher Anderson², Deepak Chandran³, Douglas Densmore⁴, Michal Galdzicki⁵, Matthew W Lux¹, Cesar A Rodriguez⁶, Guy-Bart Stan⁷ & Herbert M Sauro³

¹Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia, USA. ²Department of Bioengineering, QB3: California Institute for Quantitative Biological Research, University of California, Berkeley, California, USA.

³Department of Bioengineering, University of Washington, Seattle, Washington, USA.

⁴Department of Electrical and Computer Engineering, Boston University, Boston, Massachusetts, USA. ⁵Biomedical and Health Informatics, University of Washington, Seattle, Washington, USA. ⁶BIOFAB, Emeryville, California, USA. ⁷Department of Bioengineering and Centre for Synthetic Biology and Innovation, Imperial College London, London, UK.
e-mail: peccoud@vt.edu

1. Gardner, T.S. & Collins, J.J. US patent 6,841,376 (2005).
2. Gardner, T.S., Cantor, C.R. & Collins, J.J. *Nature* **403**, 339–342 (2000).
3. Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H. & Woicik, B. *Nature* **434**, 1125–1128 (2005).



Synthetic Biology Open Language (SBOL)

- In 2008, a small group of researchers proposed the development of the *synthetic biology open language* (SBOL), an open-source standard for the exchange of genetic designs.
- In 2011, the first version of the SBOL core data model was released. Galdzicki et al., Nature Biotechnology (2014)
- In 2013, the first version of the SBOL Visual standard was released. Quinn et al., PLoS Biology (2015)
- In 2015, SBOL Version 2.0 was officially released. Roehner et al., ACS Synthetic Biology (2016)

SBOL Community



- 100+ people from all around the world.
- 30 universities, 14 companies, 8 other types of organizations.

Organizations Supporting SBOL



EPSRC

Engineering and Physical Sciences
Research Council



Office of
Science



JOINT GENOME INSTITUTE
DEPARTMENT OF ENERGY



BioDesign
Automation
Consortium






















Current support for the development of SBOL provided by National Science Foundation Grants DBI-1356041 and DBI-1355909, and the Engineering and Physical Sciences Research Council under Grant Number EP/J02175X/1.

SBOL Governance

- **SBOL Chair** - Anil Wipat (Newcastle)
- **Past SBOL Chair** - Herbert Sauro (Washington)
- **SBOL Editors** - Bryan Bartley (Washington), Jacob Beal (Raytheon/BBN), Robert Sidney Cox (Kobe), Raik Grunberg (Montreal), and James McLaughlin (Newcastle).
- **Past SBOL Editors** - Kevin Clancy (ThermoFisher), Michal Galdzicki (Washington), Goksel Misirli (Newcastle), Ernst Oberortner (DOE JGI), Matthew Pocock (Newcastle), Jacqueline Quinn (Google), Cesar Rodriguez (Autodesk), Nicholas Roehner (Boston), and Mandy Wilson (VBI).
- **Steering Committee** - Jacob Beal (Raytheon/BBN), Kevin Clancy (ThermoFischer), Douglas Densmore (Boston), John Gennari (Washington), Nathan Hillson (JBEI), Chris Myers (Utah), and Herbert Sauro (Washington).
- **Advisory Panel** - Matthew Chang (Singapore), Victor de Lorenzo (CSIC), Traci Haddock (iGEM), Richard Kitney (Imperial), Sarah Munro (NIST), and Chris Voigt (MIT).

SBOL Visual (Version 1.0)

 promoter	 origin of replication
 cds	 primer binding site
 ribosome entry site	 blunt restriction site
 terminator	 sticky restriction site
 operator	 5' overhang
 insulator	 3' overhang
 ribonuclease site	 assembly scar
 rna stability element	 signature
 protease site	 user defined
 protein stability element	

New symbols
added on
community
consensus.

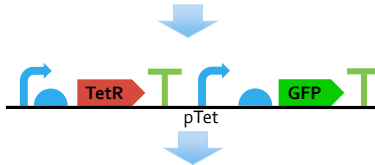
Quinn et al., PLoS Biology (2015)

SBOL Data Model (Version 2.0)

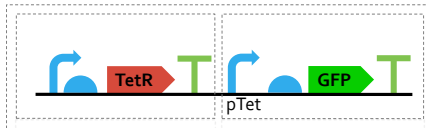
FASTA

ACTGTGCCGTTAAACGTGATTAAATCCGTACTGATAT...

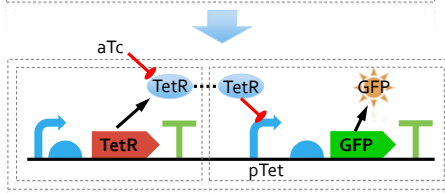
GenBank



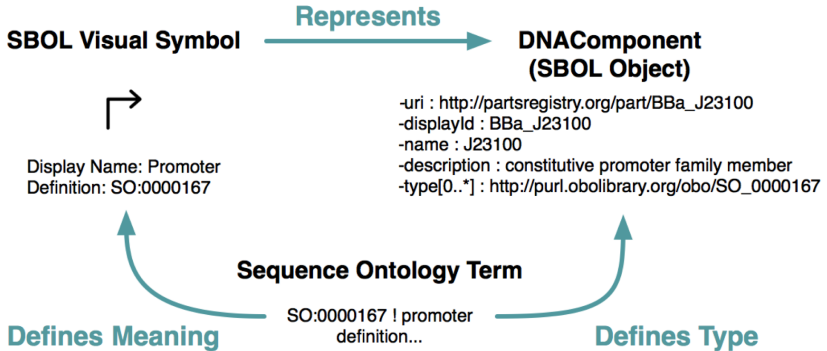
SBOL 1.1



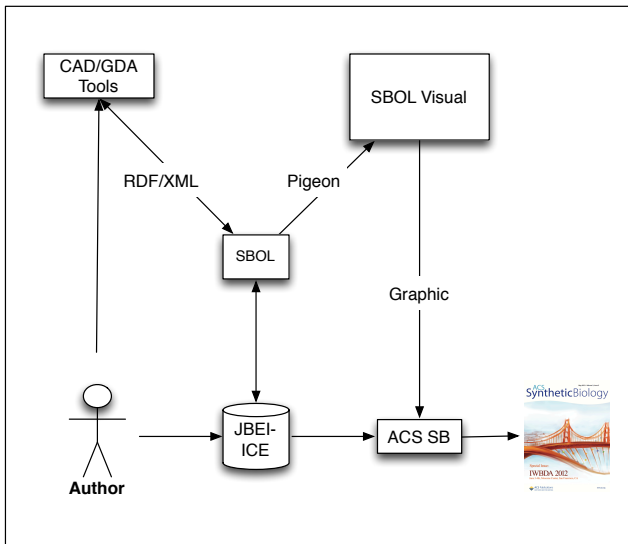
SBOL 2.0



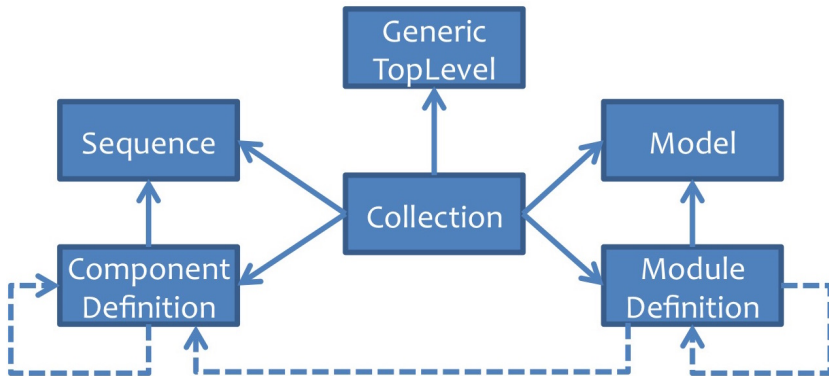
Connection Between Visual and Data Model



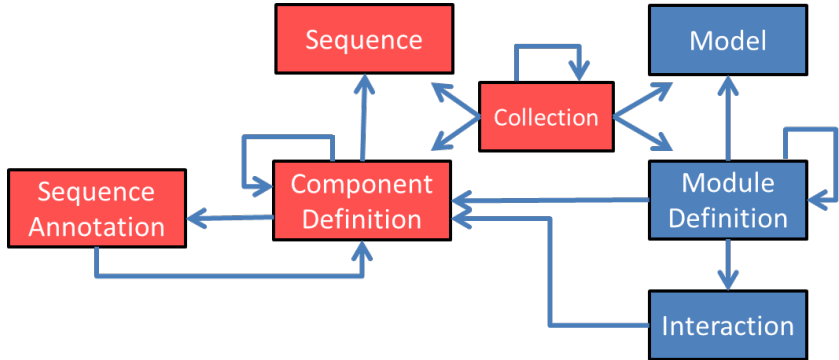
SBOL and ACS Synthetic Biology



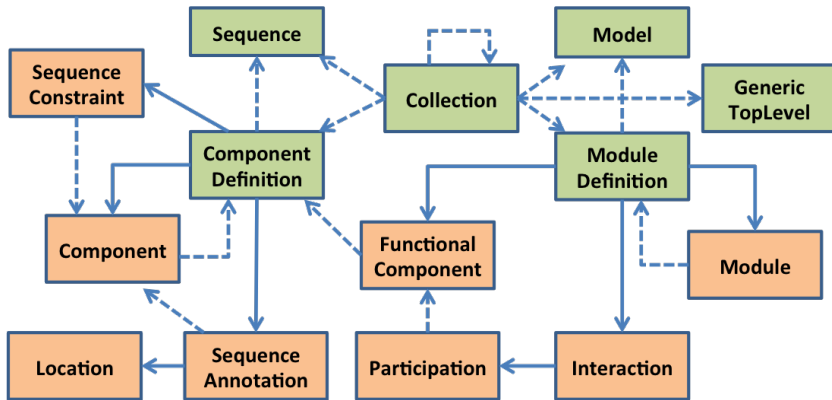
SBOL 2.0 Data Model



SBOL 2.0 Data Model



SBOL 2.0 Data Model



Annotations

```
<?xml version="1.0" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:myerslab="http://www.async.ece.utah.edu"
  xmlns:sbol="http://sbols.org/v2#">
  <sbol:Collection rdf:about="http://www.async.ece.utah.edu/My_Parts/1.0">
    <dc:title>My_Parts</dc:title>
    <dc:description>These_are_my_parts.</dc:description>
    <sbol:member rdf:resource="http://www.async.ece.utah.edu/pLac/1.0"/>
    ...
  </sbol:Collection>
  <sbol:ModuleDefinition rdf:about="http://www.async.ece.utah.edu/LacI_Inverter/1.0">
    <sbol:role rdf:resource="http://www.async.ece.utah.edu/inverter"/>
    <sbol:component>
      <sbol:FunctionalComponent rdf:about="http://www.async.ece.utah.edu/LacI_Inverter/LacI/1.0">
        <sbol:definition rdf:resource="http://www.async.ece.utah.edu/LacI/1.0"/>
        <sbol:access rdf:resource="http://sbols.org/v2#public"/>
        <sbol:direction rdf:resource="http://sbols.org/v2#input"/>
      </sbol:FunctionalComponent>
    </sbol:component>
    ...
  </sbol:ModuleDefinition>
  ...
</rdf:RDF>
```


Annotations (cont)

```
<sbol:interaction>
  <sbol:Interaction
    rdf:about="http://www.async.ece.utah.edu/LacI_Inverter/pLac_Repression/1.0">
    <sbol:type rdf:resource="http://identifiers.org/sbo/SBO:0000169"/>
    <myerslab:parameter rdf:resource="http://www.async.ece.utah.edu/Kr_LacI">
    <sbol:participation>
      <sbol:Participation
        rdf:about="http://www.async.ece.utah.edu/LacI_Inverter/pLac_Repression/p1/1.0">
        <sbol:role rdf:resource="http://www.async.ece.utah.edu/repressed"/>
        <sbol:participant
          rdf:resource="http://www.async.ece.utah.edu/LacI_Inverter/pLac/1.0"/>
        </sbol:Participation>
      </sbol:participation>
    <sbol:participation>
      <sbol:Participation
        rdf:about="http://www.async.ece.utah.edu/LacI_Inverter/pLac_Repression/p2/1.0">
        <sbol:role rdf:resource="http://identifiers.org/sbo/SBO:0000020"/>
        <sbol:participant
          rdf:resource="http://www.async.ece.utah.edu/LacI_Inverter/LacI/1.0"/>
        </sbol:Participation>
      </sbol:participation>
    </sbol:Interaction>
  </sbol:interaction>
  ...
</sbol:ModuleDefinition>
<myerslab:Parameter rdf:about="http://www.async.ece.utah.edu/Kr_LacI">
  <myerslab:type rdf:resource="http://identifiers.org/sbo/SBO:0000281"/>
  <myerslab:value>0.5</myerslab:value>
</myerslab:Parameter>
...
</rdf:RDF>
```

Library Support for SBOL 2.0

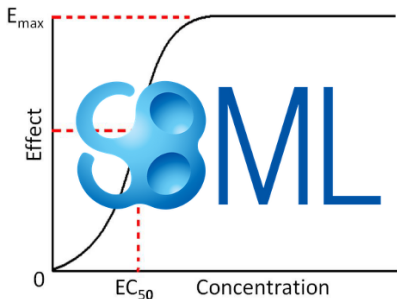
- Crucial to the success of a standard is software infrastructure to support developers' integration of the standard within their tools.
- There are several library implementations of the SBOL data structure, which provide an *application programmers interface* (API) for tool developers to interact with SBOL data objects.
 - libSBOLj - native Java library
 - libSBOL - C/C++ library
 - pySBOL - Python library
 - sboljs - Javascript library
- Library distributions include detailed documentation for the class definitions and the methods provided by the API.
- An online validator/converter powered by libSBOLj is available from the SBOL website.

Partial List of SBOL Compliant Software Tools

Name	Repository	DNA	Design	Visual	Import	Export	Lossless
Benchling		x				1.1	
Cello		x	x	x		2.0	
DeviceEditor		x		x	1.1	1.1	
DNAPlotLib		x		x	1.1	1.1	
Eugene		x			1.1	1.1	
GeneGenie		x				2.0	
GenoCAD		x		x		1.1	
GraphViz		x		x			
GSL/Thumper		x			2.0	2.0	
iBioSim			x		2.0/1.1	2.0	x
ICE	x			x	2.0	2.0	x
j5		x			1.1	1.1	
Kera		x		x			
MoSeC			x			1.1	
Pigeon		x		x			
Proto BioCompiler			x	x		1.1	
Raven		x		x			
SBOL Stack	x				2.0	2.0	x
SBROME			x		1.1	1.1	
TeselaGen		x		x	1.1	1.1	
TinkerCell		x	x	x	1.1	1.1	
Vector Editor		x			1.1	1.1	
VirtualParts	x				2.0	2.0	x
VisBOL		x	x	x	2.0/1.1		
SBOL Designer		x		x	1.1	1.1	x

WARNING:
Changing rapidly,
so likely not
completely
accurate.

Connecting SBML to SBOL



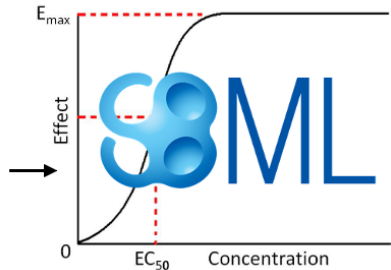
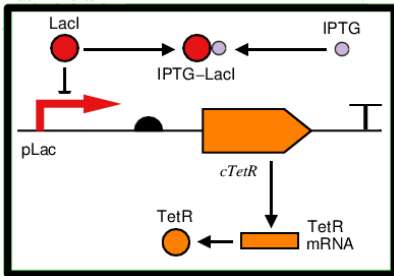
RDF/XML
Annotation



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

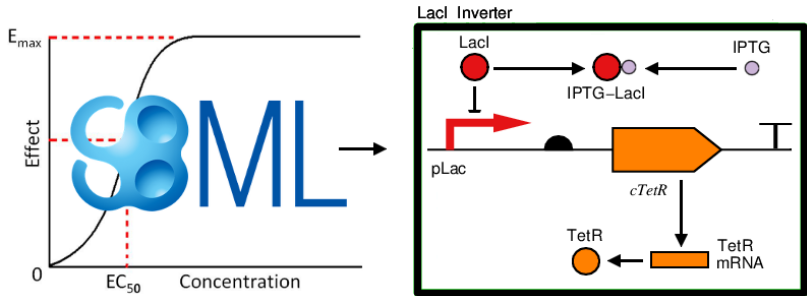
Converting SBOL to SBML

LacI Inverter



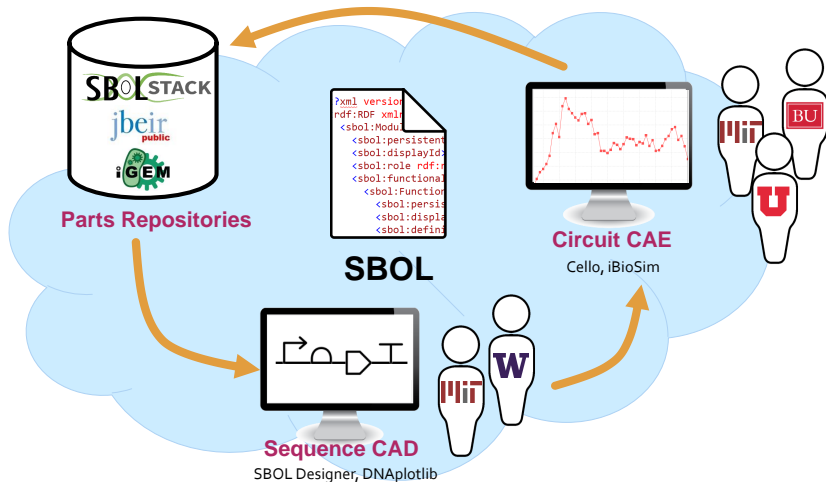
Roehner et al., ACS Synthetic Biology (2014)

Converting SBML to SBOL

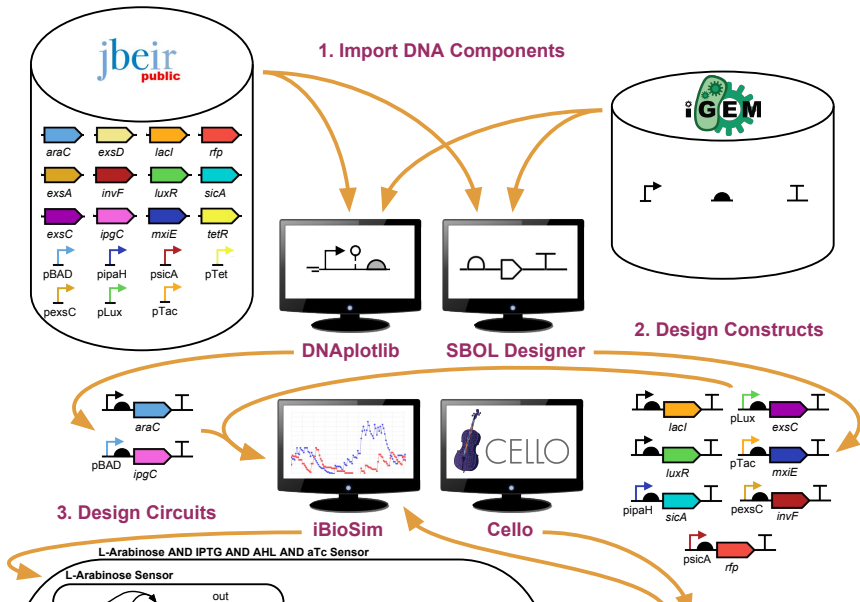


Nguyen et al., ACS Synthetic Biology (2016)

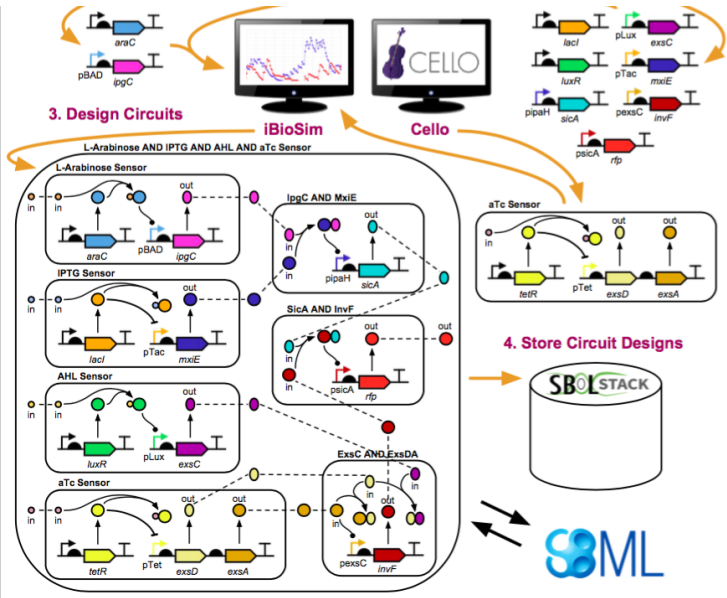
Synthetic Biology Workflow Using SBML/SBOL



Synthetic Biology Workflow Using SBML/SBOL



Synthetic Biology Workflow Using SBML/SBOL



Conclusion

- Standards are an important enabler for data sharing and reproducibility in systems and synthetic biology.
- SBOL is supported by libraries that provide an API, documentation, and utilities for managing the encoding and exchange of data.

More Information

- More information about SBOL: <http://www.sbolstandard.org/>.
- All supporting libraries are open source.
- Please join the SBOL communities!