Synthetic Biology Open Language

Presenter: Jacob Beal (BBN Technologies)
Introductory SBOL Workshop, August 1st, 2018
Berkeley, CA

Editors: Angel Goni-Moreno (Newcastle University), Curtis Madsen (Boston University), Umesh P (University of Kerala), Zach Palchick (Zymergen), Nicholas Roehner (BBN Technologies)

Chair: Anil Wipat (Newcastle University)

SBOL Development Group: ~120 members from more than 60 academic and industrial organizations.

- Three foundational principles of synthetic biology based on engineering practice (Endy 2005):
  - **Standardization**
  - **Abstraction**
  - **Decoupling**
- Synthetic biology was born with the broad goal of engineering or 'wiring' biological circuitry — be it genetic, protein, viral, pathway or genomic — for manifesting logical forms of cellular control. (Collins 2010)
- Biology has long surpassed its mainly descriptive stage, and the questions now asked are increasingly amenable to experimental approaches and theoretical concepts taken from the physical and engineering sciences. (Scwhille 2011)

Standards are a foundational principle of synthetic biology

What is SBOL?

SBOL actually consists of 2 standards:

- **SBOL Data model**
  - a formalized representation of data objects
- **SBOL Visual**
  - a standardized set of schematic symbols for genetic design

To guarantee interoperability between tools and standards, SBOL leverages **ontologies**, such as the Sequence Ontology (SO), Systems Biology Ontology (SBO).

Milestones in the history of SBOL

Apr, 2008	**Kick-off** at a computational synthetic biology workshop at the University of Washington
Jun, 2011	The **SBOL Developers Group** was officially established with adoption of formal rules of governance and election of editors
Sep, 2011	First version of the **SBOL data model** was released (BB FRFC 84)
Mar, 2013	First version of the **SBOL Visual** standard was released (BBF RFC 93)
Jul, 2015	**SBOL Version 2.0** was officially released (BBF RFC 108)
Jun, 2016	**ACS Synthetic Biology adopts SBOL** as publication standard
Dec, 2017	**SBOL Visual 2.0** officially released **SBOL Version 2.2** supports complete design-build-test cycle

Evolution of standards for Bioinformatics

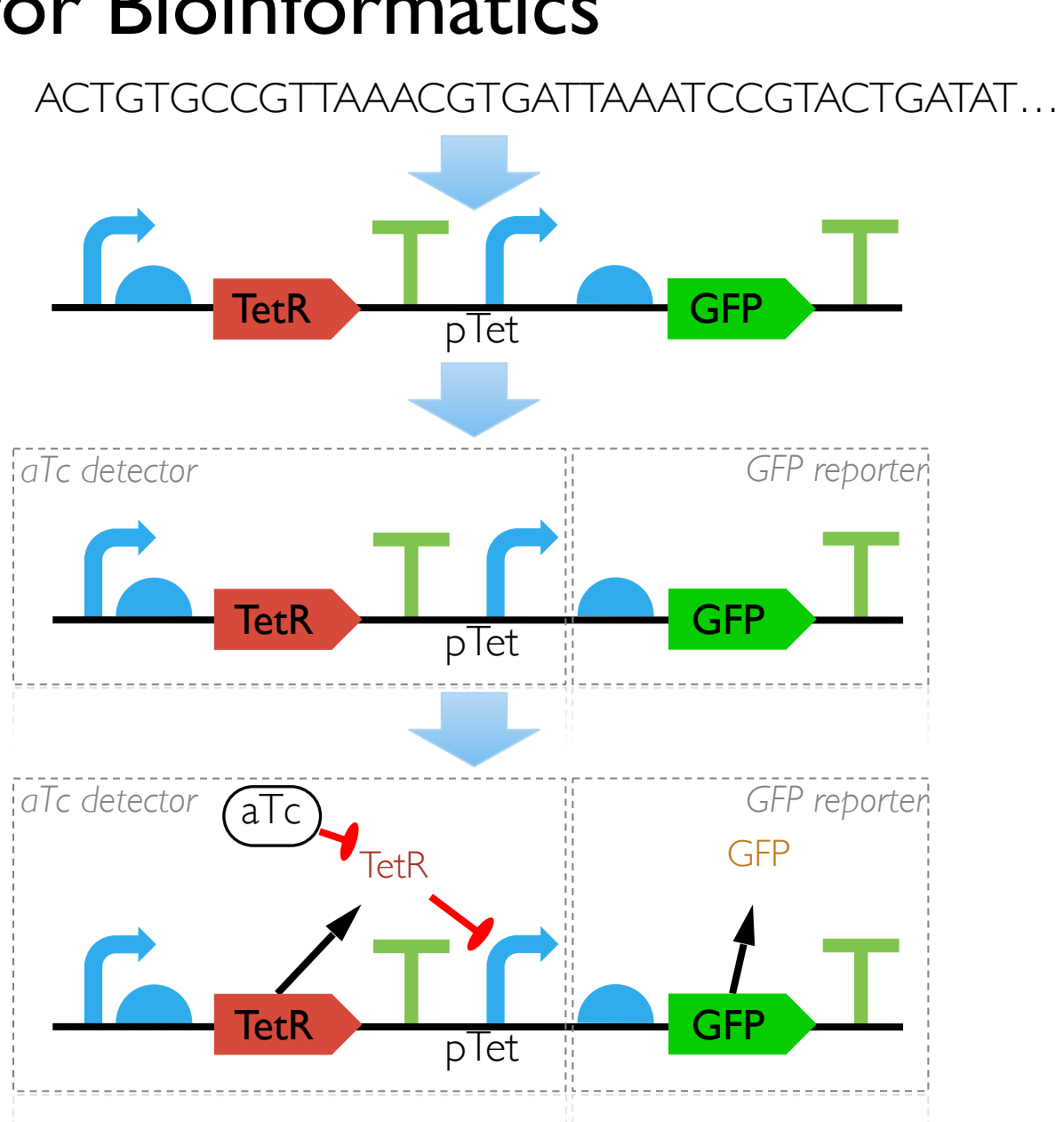
FASTA
Raw sequence information

GenBank
Contains sequence features

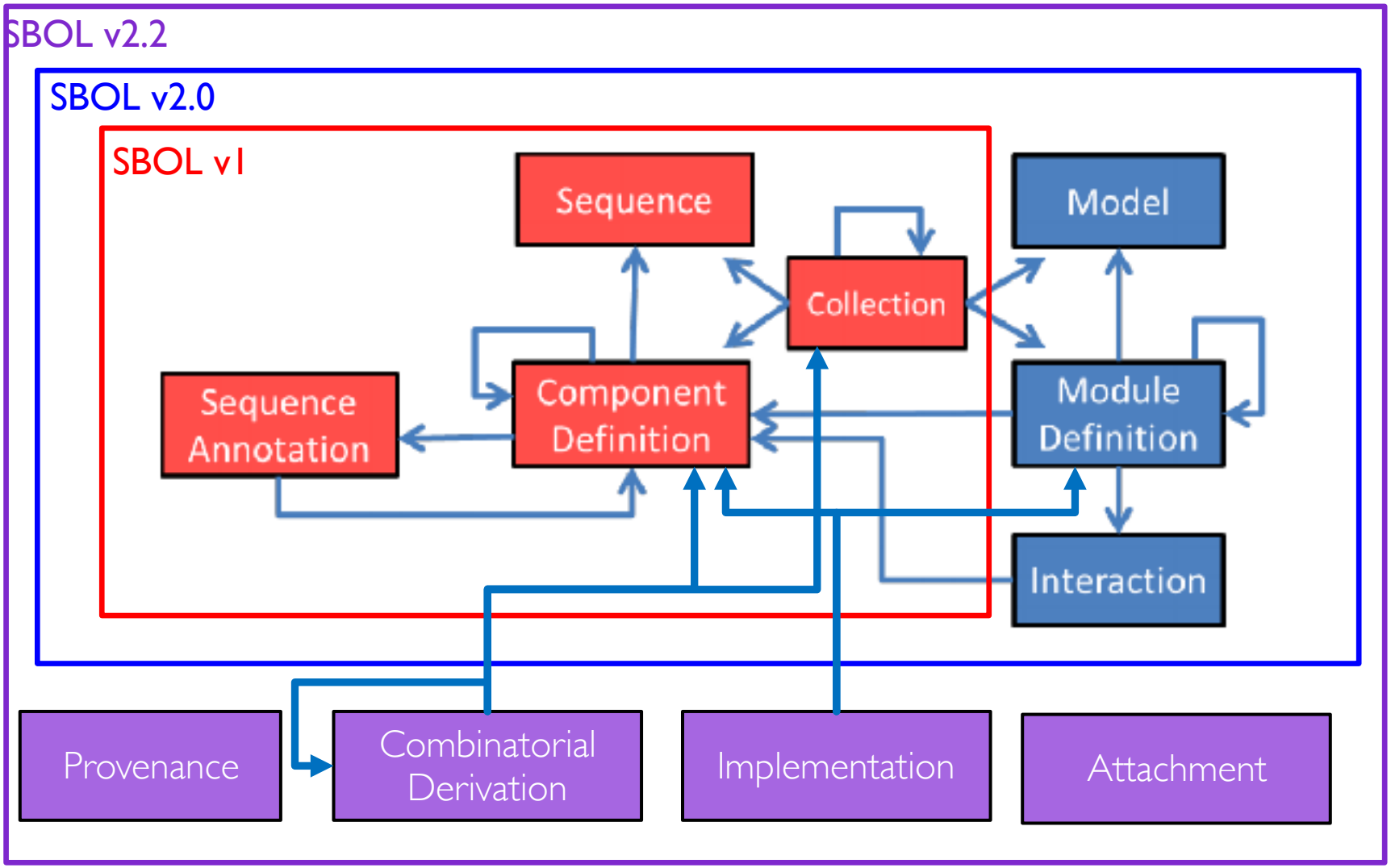
SBOL 1
Represents sequence features hierarchically

SBOL 2

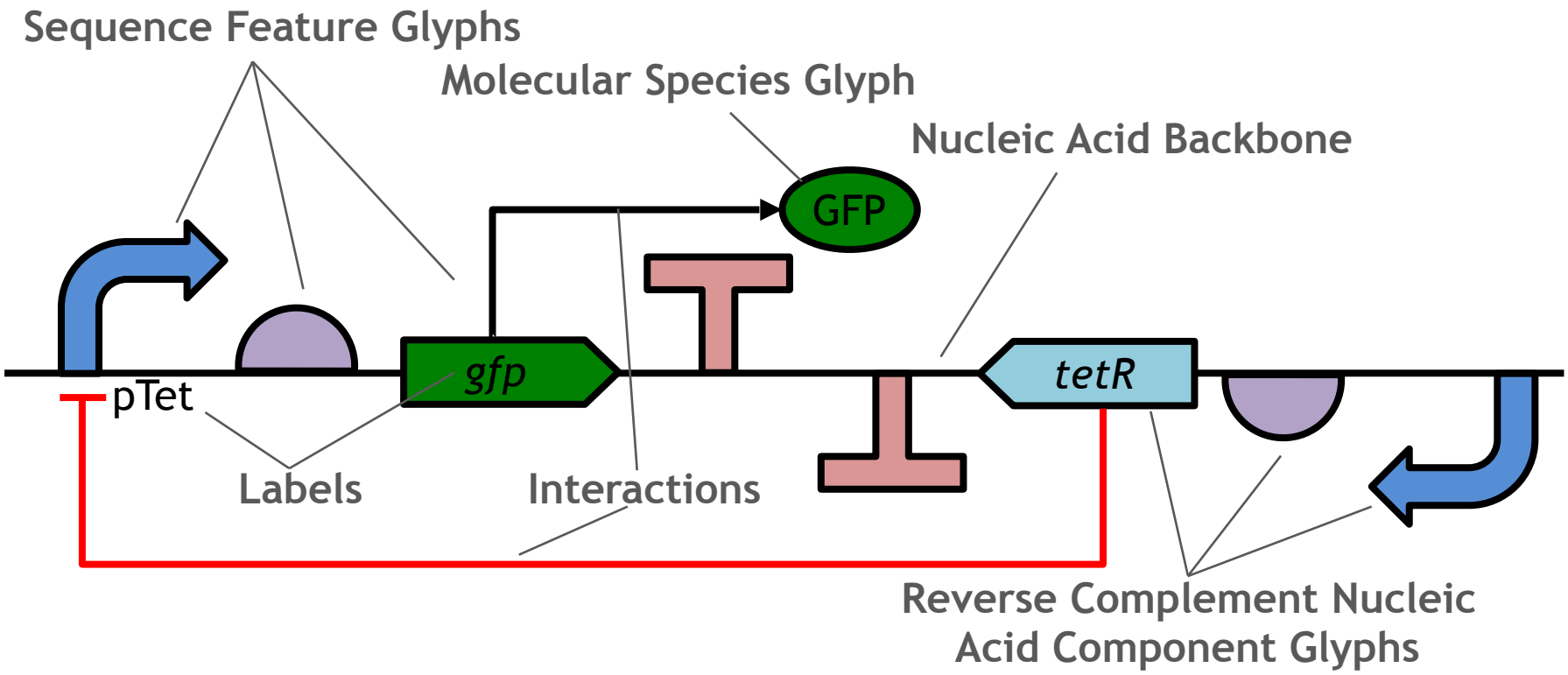
- Represents additional molecule types
- Represents modules with inputs and outputs



SBOL Data Model

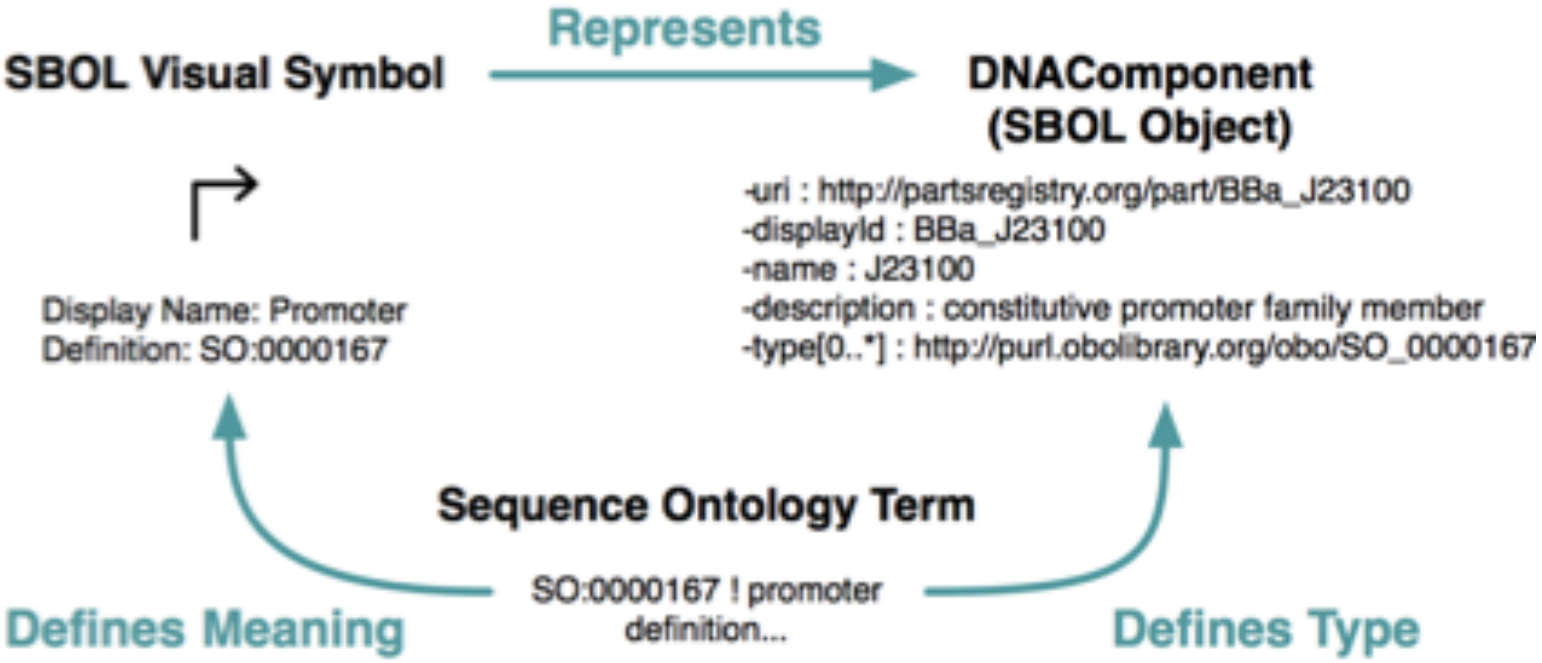


SBOL Visual v2.0

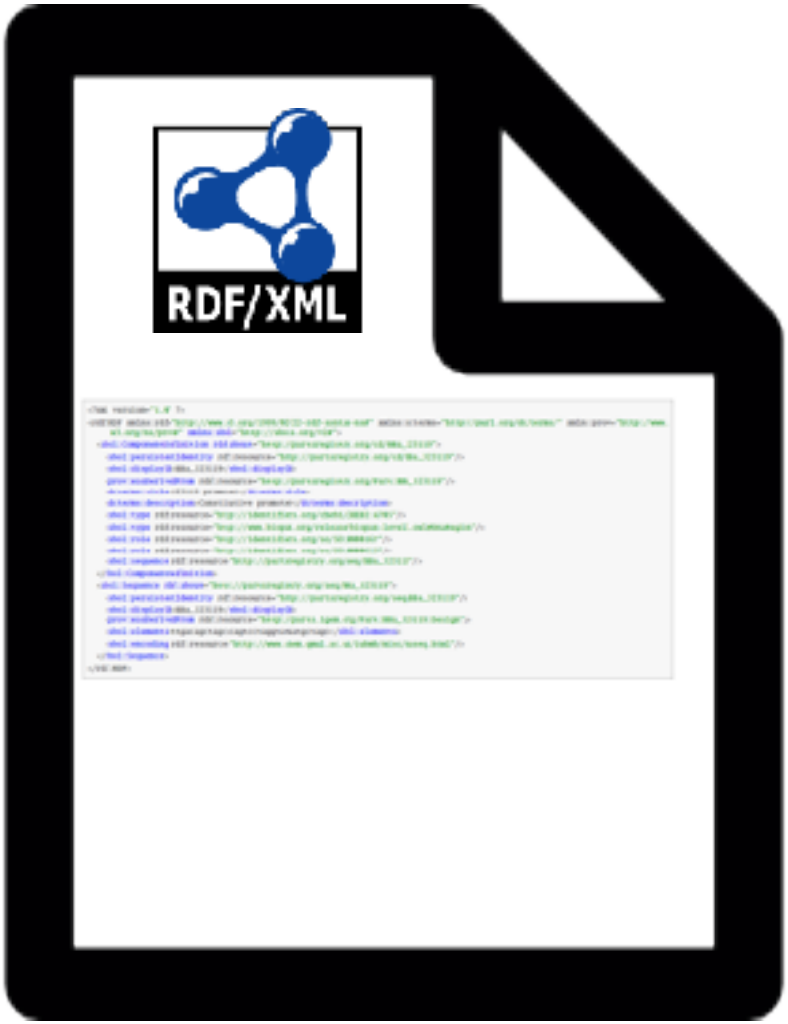


Grey text and lines (including this) are annotations

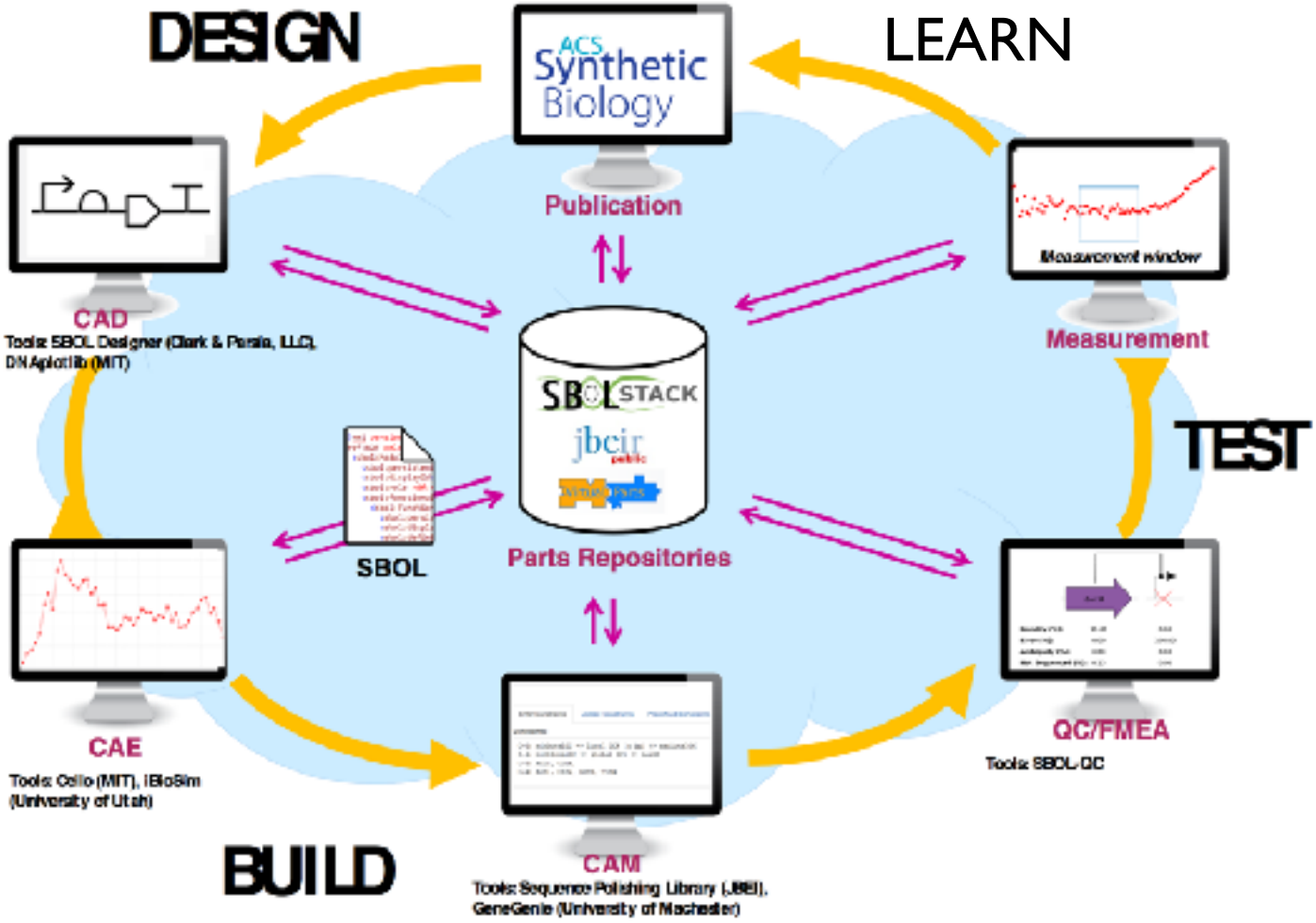
Relation between SBOL Visual and SBOL Data Model



SBOL is serialized in RDF/XML Format

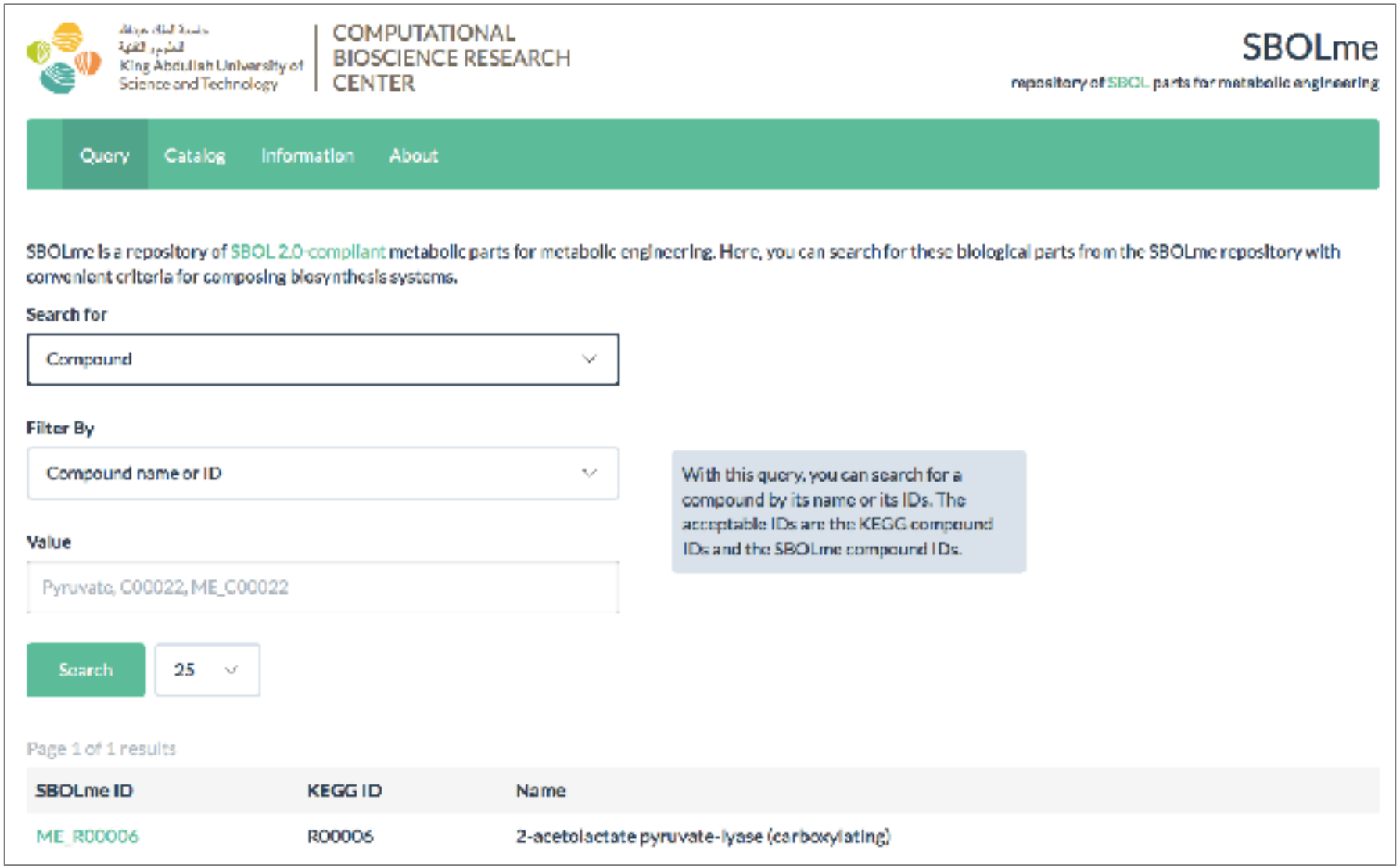


- XML allows data to be structured in hierarchical trees, is well-supported and well-understood by software developers
- RDF makes data integration across networks easier



SBOL helps synthetic biologists to collaborate across different stages of (automation-assisted) workflows

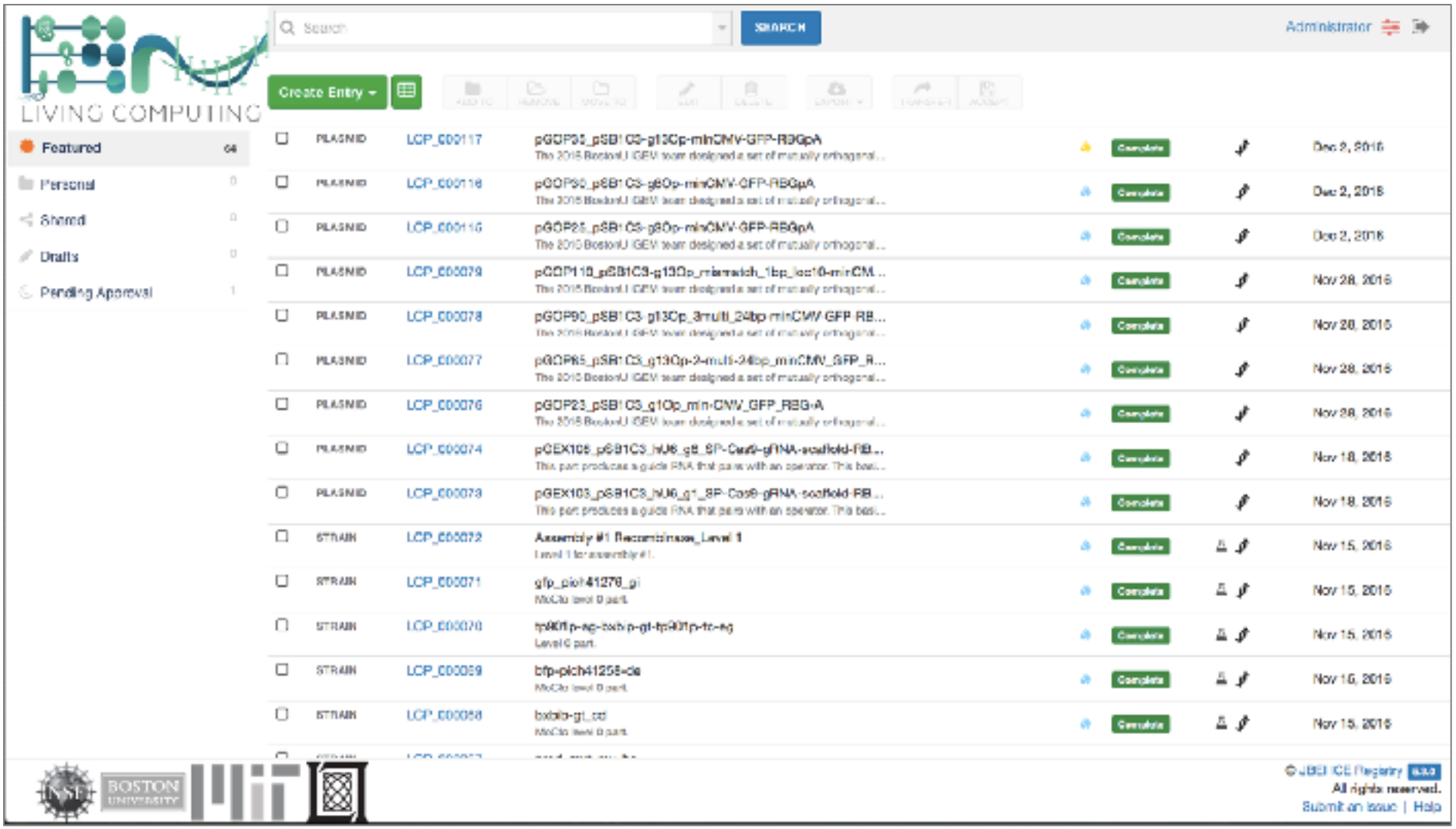
Data Repositories: SBOLme



Kuwahara et al., ACS Synthetic Biology (2017)

Includes 28,437 chemical compounds, 6,883 enzyme classes, 9,909 metabolic reactions, and 3,173,238 proteins from 3,908 organisms.

Data Repositories: ICE



Ham et al., Nucleic Acid Research (2012)

Data Repositories: SynBioHub

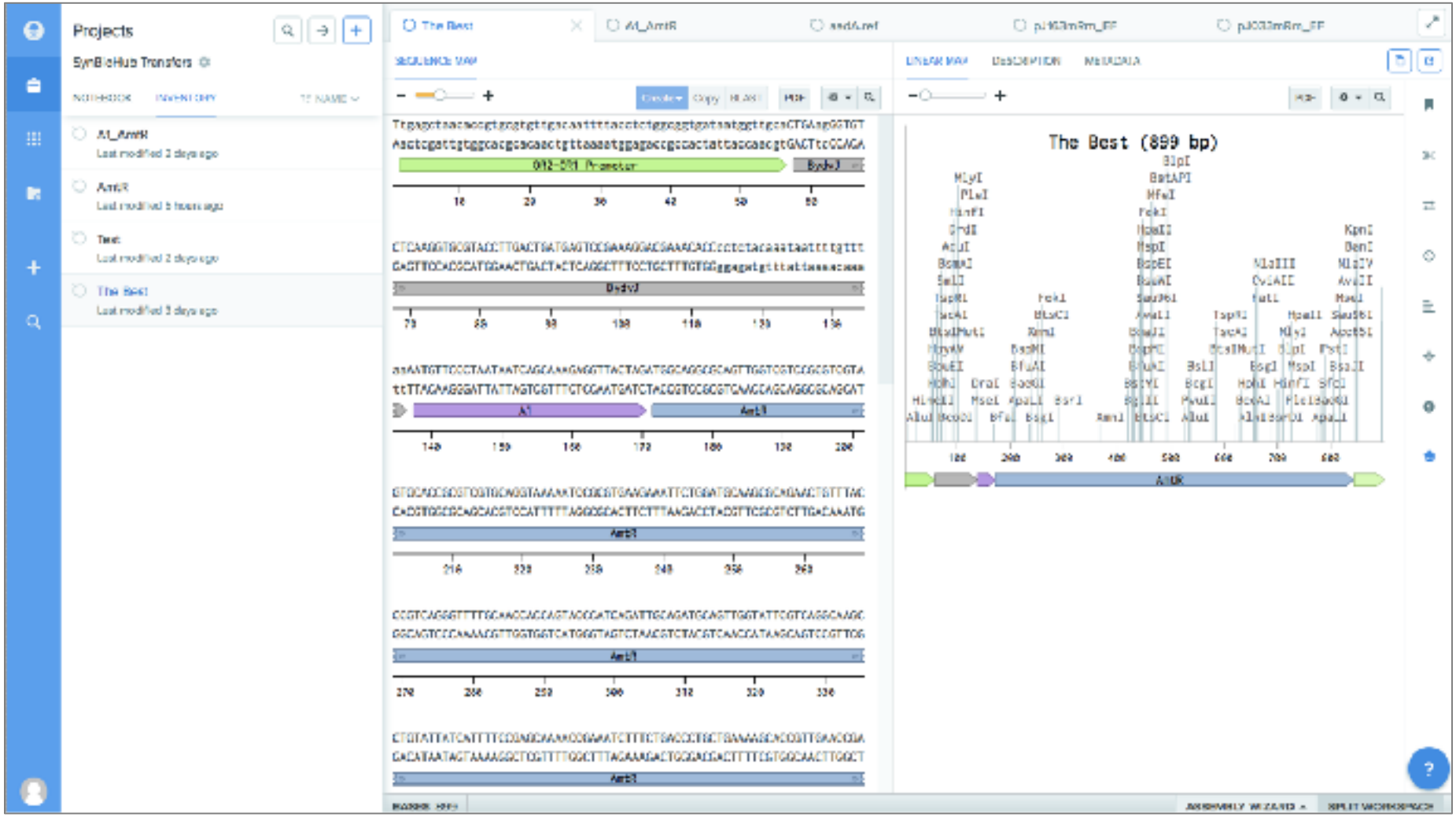
Newcastle University

James McLaughlin
Anil Wipat

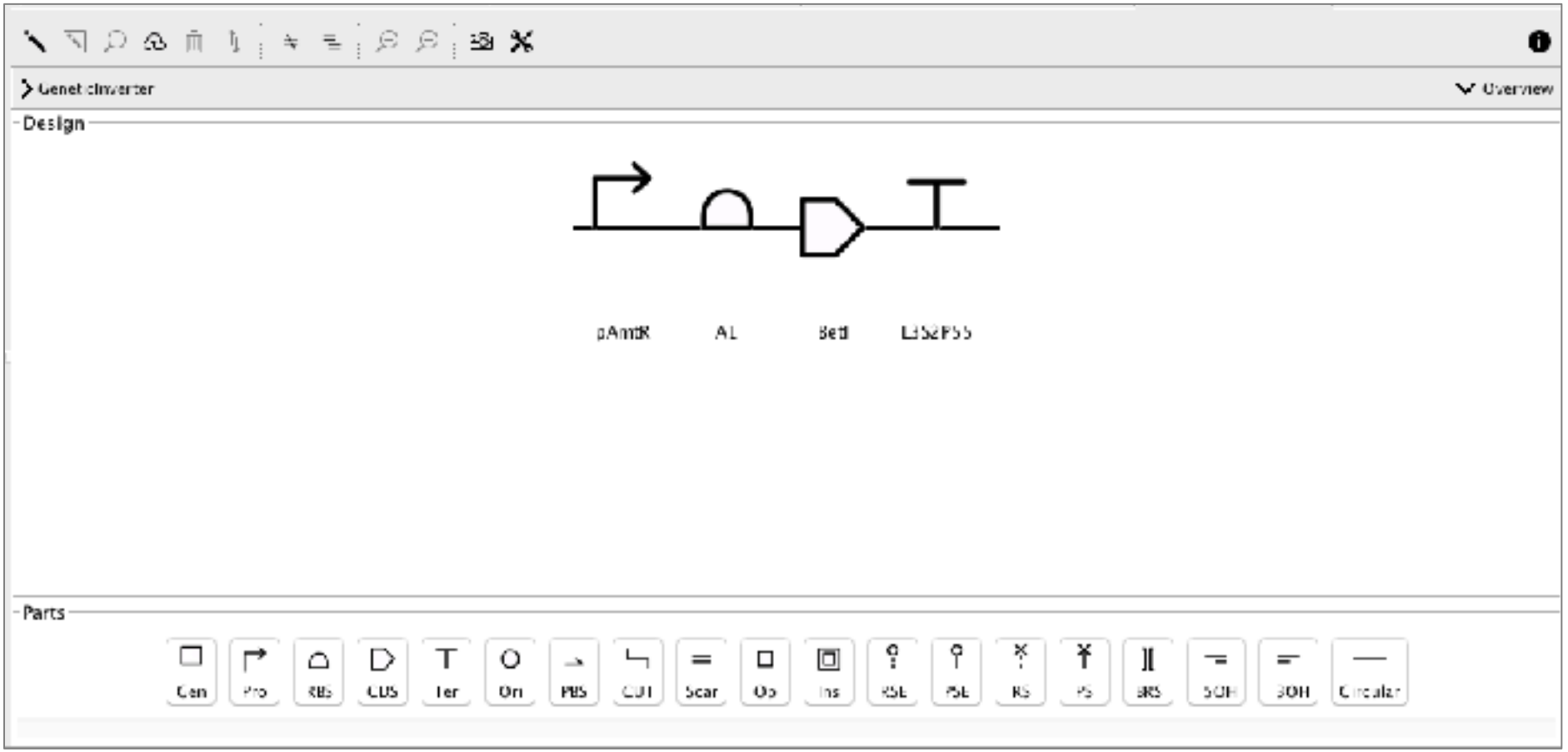
THE UNIVERSITY OF UTAH

Zach Zundel
Chris Myers

Sequence Editor: Benchling



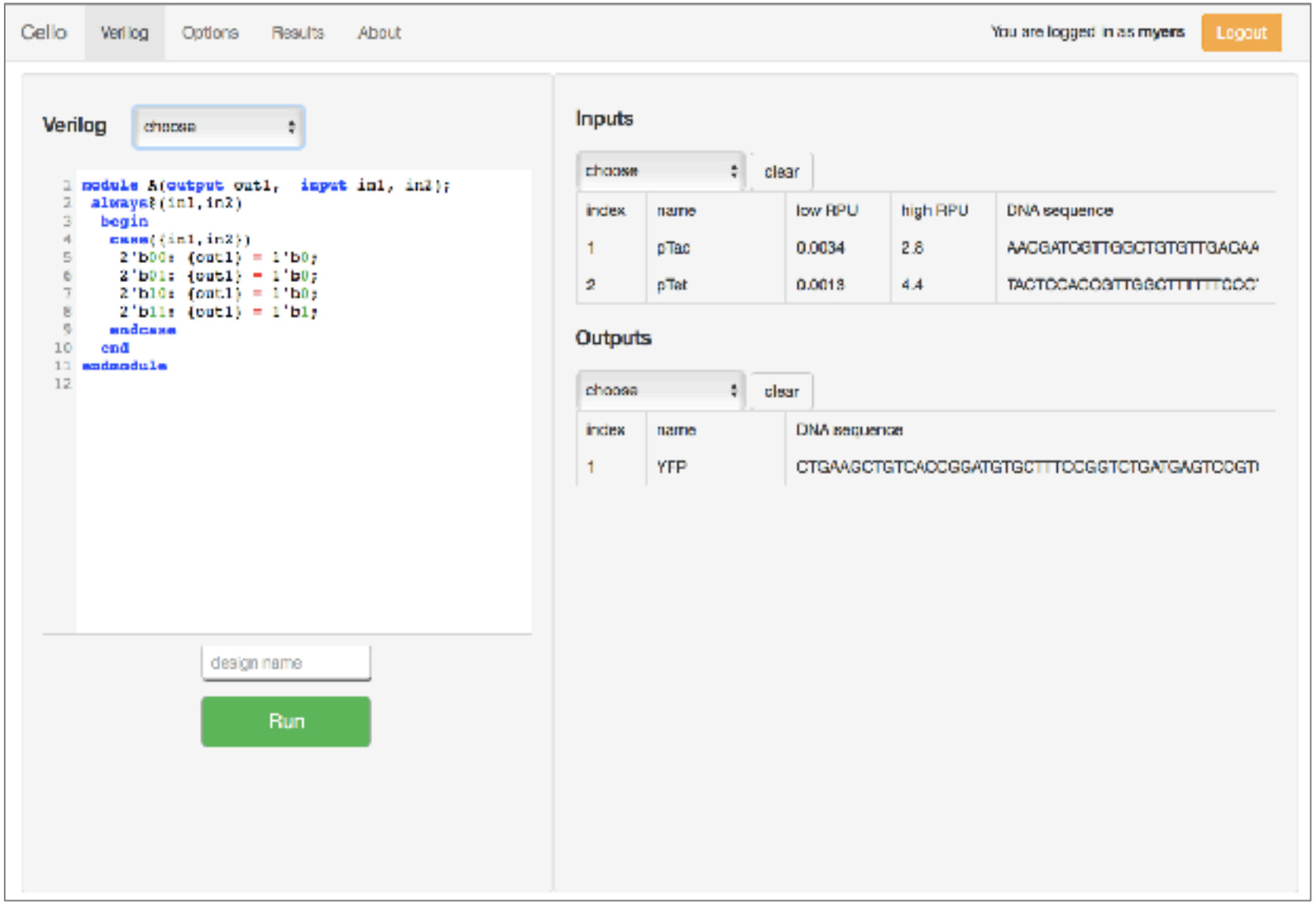
Sequence Editor: SBOL Designer



Zhang et al., ACS Synthetic Biology (2017)

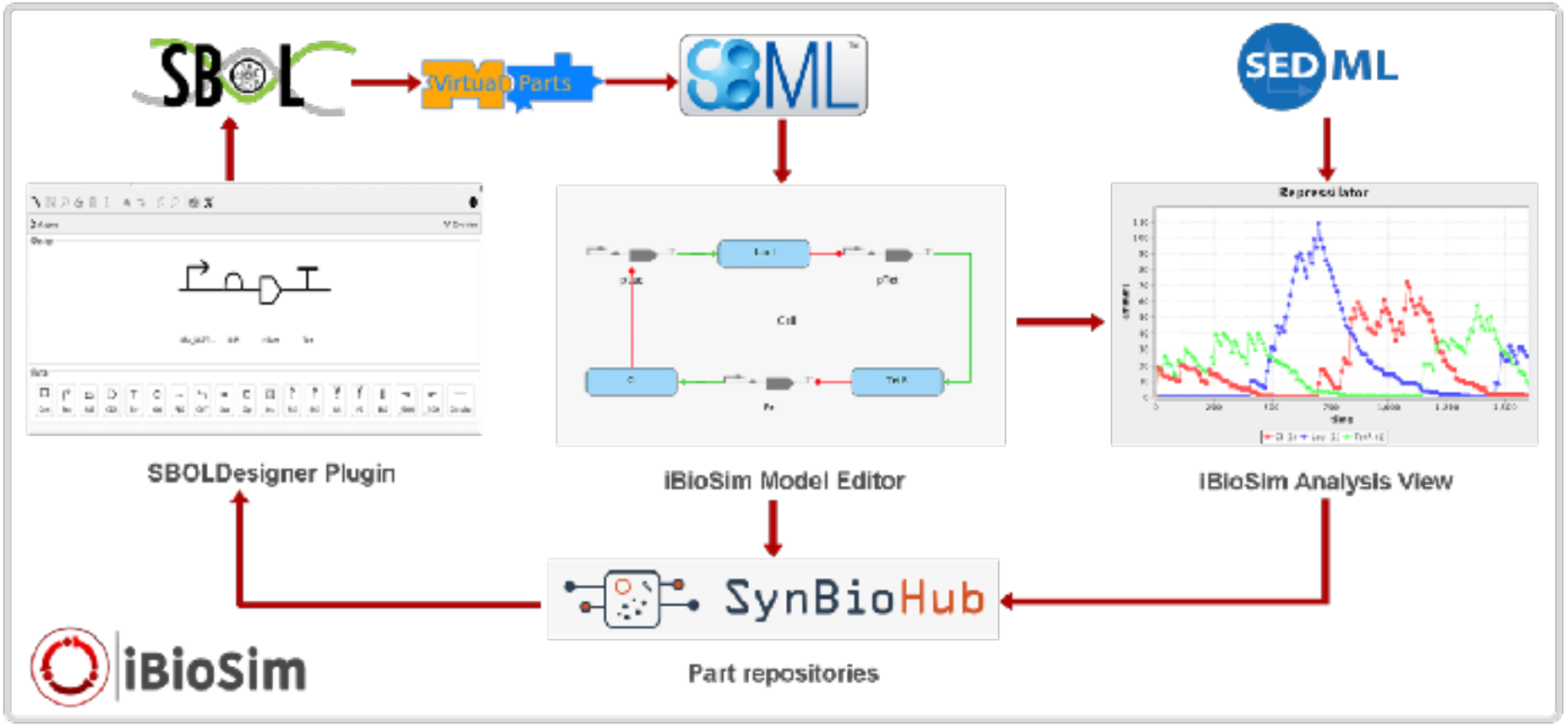
Other sequence editors that support SBOL: **DeviceEditor, J5, VectorEditor (JBEI), DNAPlotLib (MIT/UW/Bristol), Eugene (Boston), GenoCAD (VBI), BOOST (JGI), etc.**

Circuit Design Tools: Cello



Nielsen et al., Science (2016)

Circuit Design Tools: iBioSim



Madsen et al., IEEE Design & Test (2012)

Open Source Libraries

Software libraries which import and export SBOL files are freely available at the Synthetic Biology Data Exchange on GitHub under the Apache 2.0 license. Libraries are implemented in:

- Java
- Javascript
- C/C++
- Python

Developer support includes (see http://sbolstandard.org/software/libsbol)

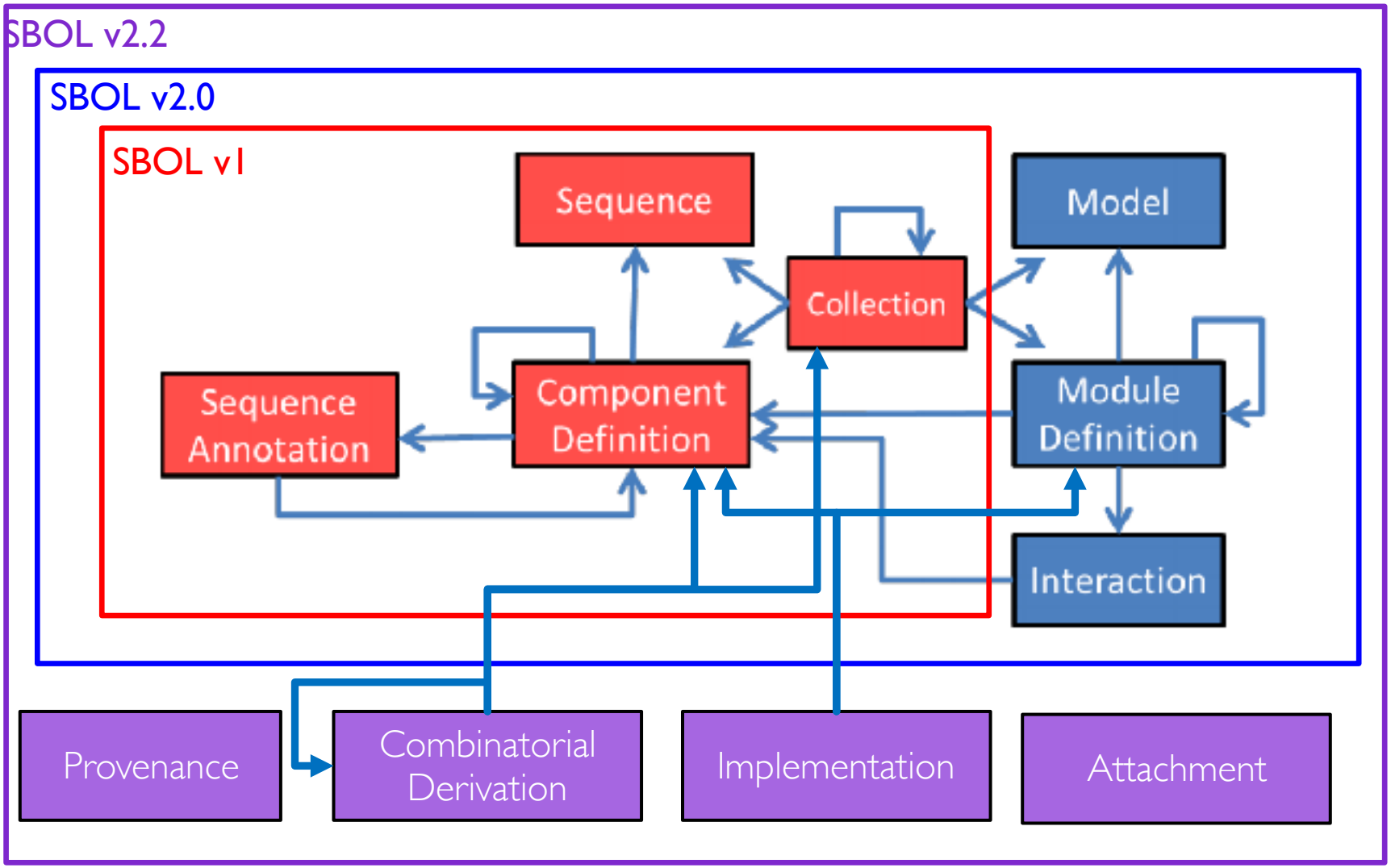
- Online documentation
- Getting started tutorials
- Sample projects
- Code examples

ACS Synthetic Biology has officially adopted SBOL as publication standard

- **SBOL Visual** is the recommended graphical notation for depicting genetic constructs
- **SBOL 2.0 Data Model** is the preferred format for nucleic acid sequences.
- Manuscript submission, review, and production process is linked to SBOL-enabled repositories
- **Joint Bioenergy Institute (JBEI)** has set up the initial repository.

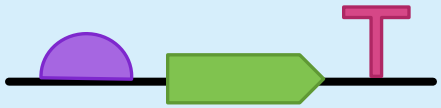
Read all about it in the ACS Synthetic Biology viewpoint article: Improving Synthetic Biology Communication: Recommended Practices for Visual Depiction and Digital Submission of Genetic Designs, or listen to the interview with Jake Beal and Nathan Hillson.

SBOL Data Model

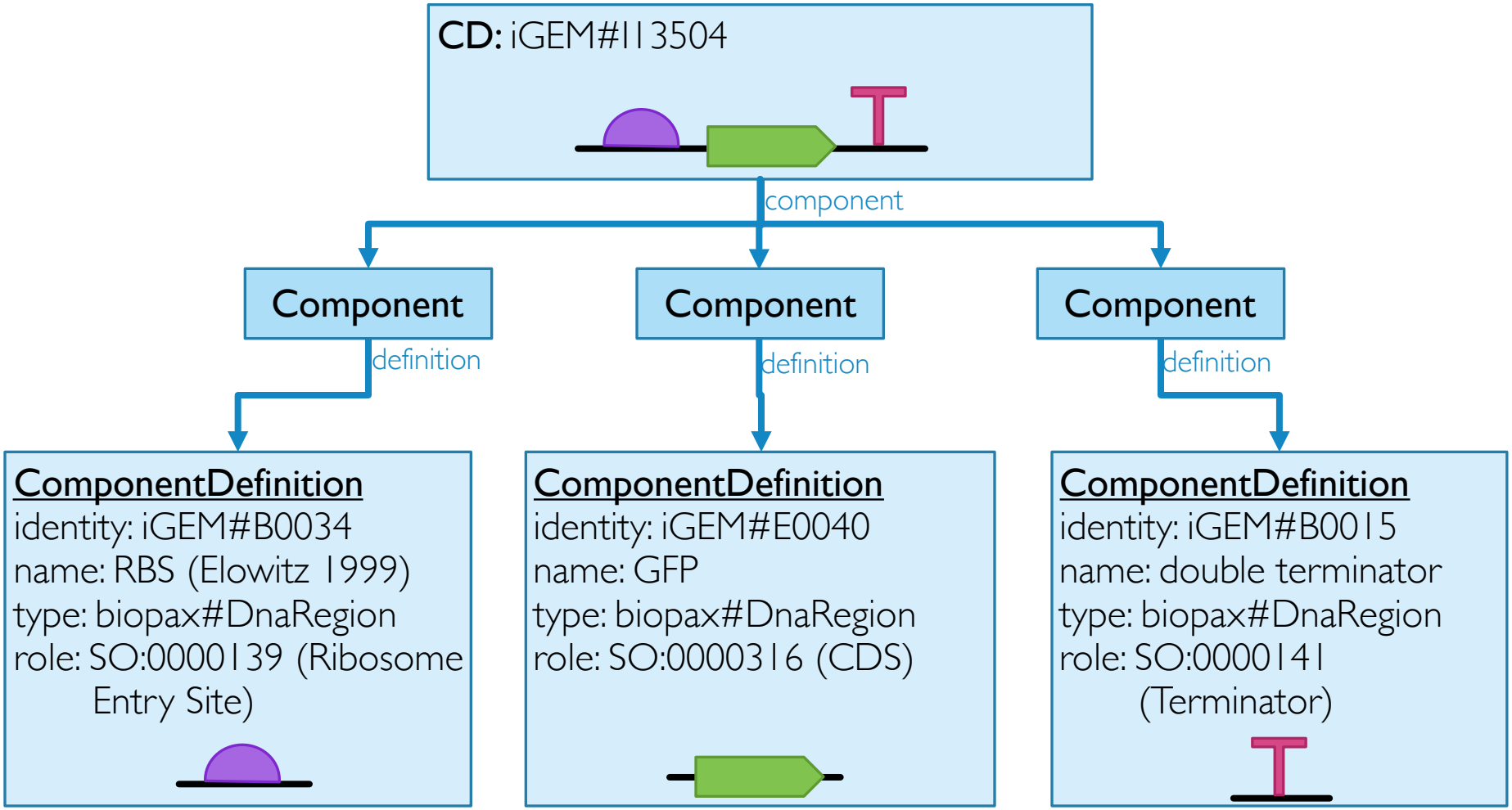


ComponentDefinition

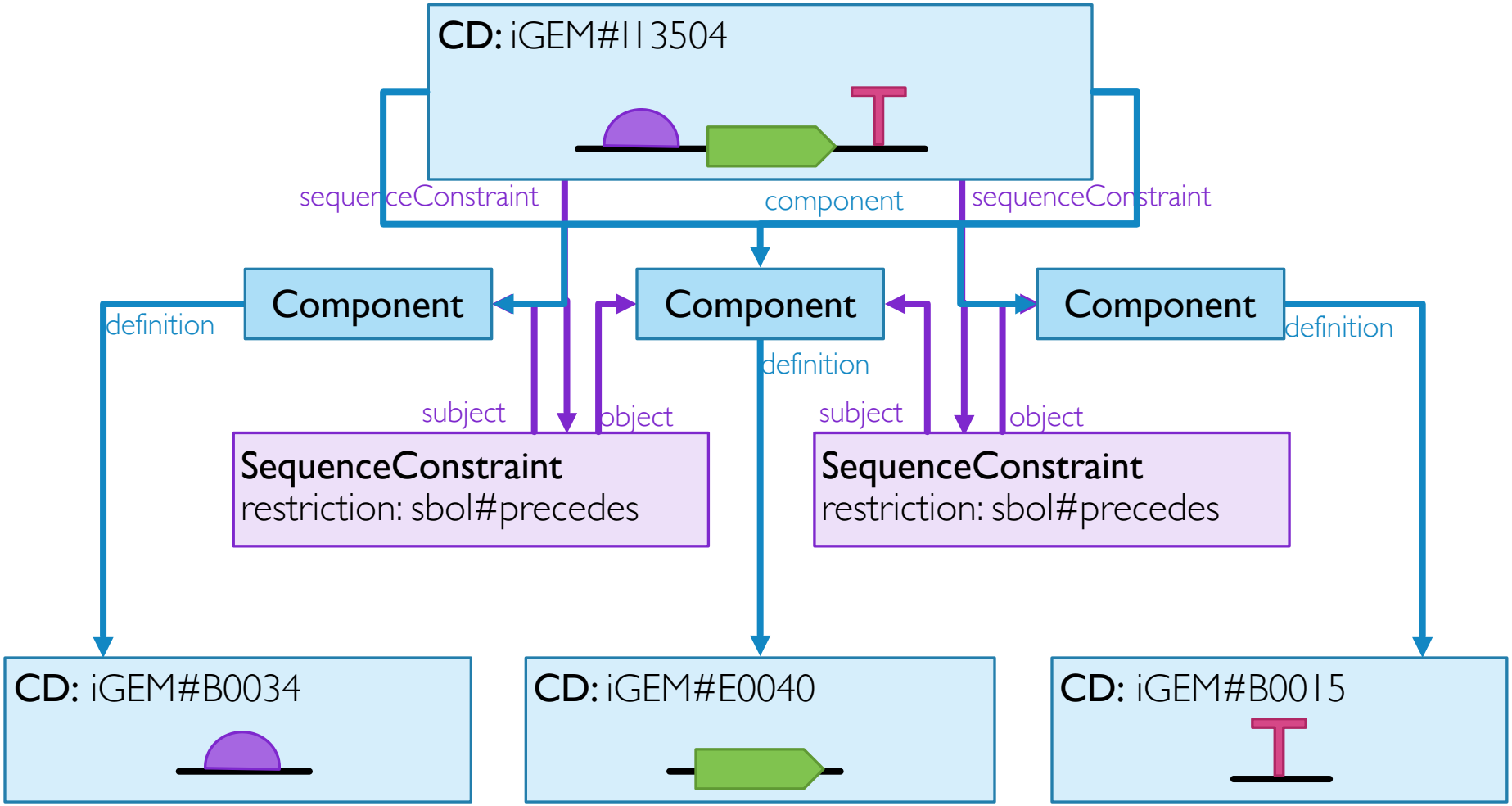
ComponentDefinition
identity: iGEM#113504
name: "iGEM 2016 interlab reporter"
description: "GFP expression cassette
used for 2016 iGEM interlab"
type: biopax#DnaRegion
role: SO:0000804 (Engineered Region)



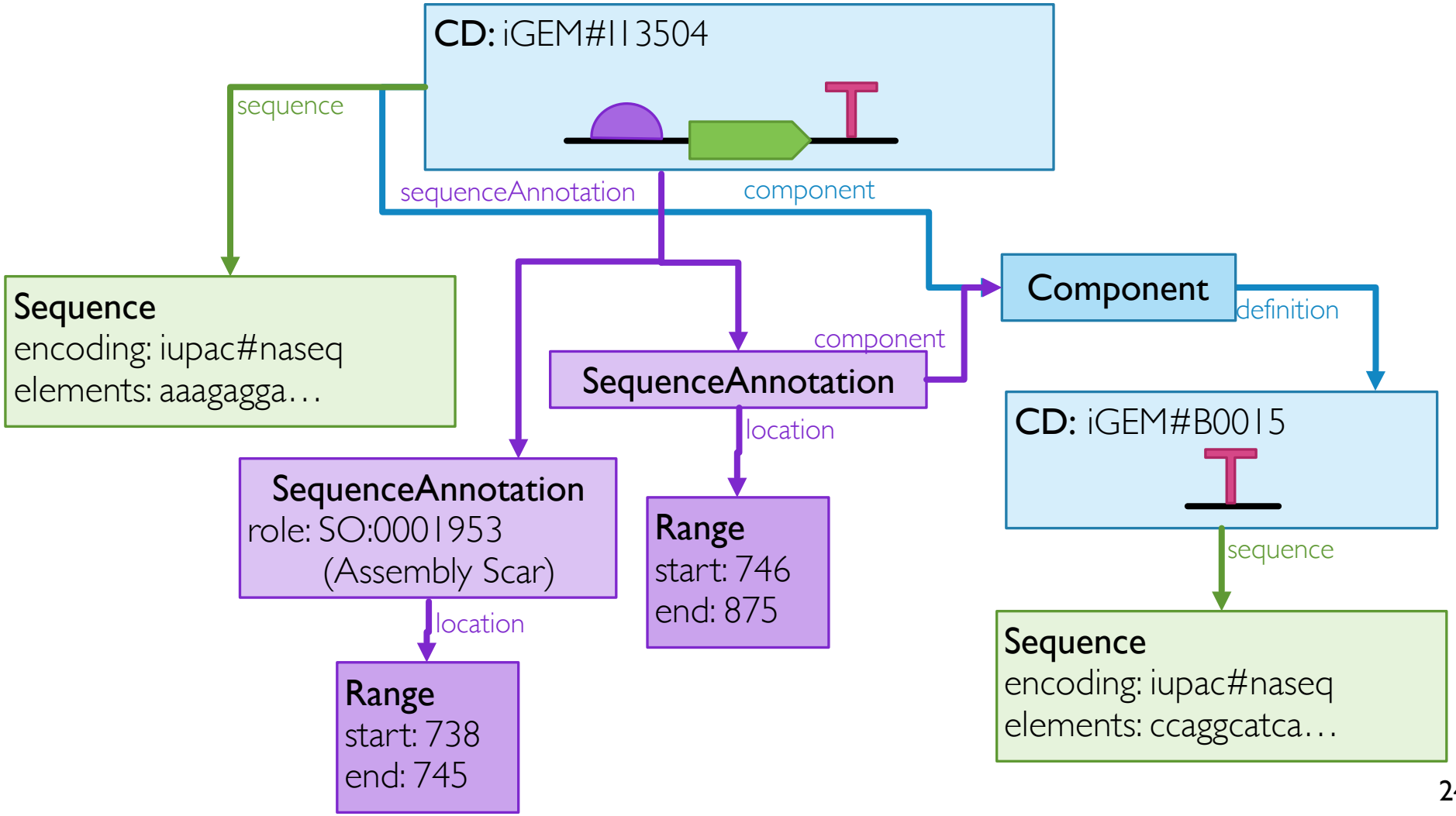
Component


CD: iGEM#113504
component
Component
Component
Component
definition
definition
definition
ComponentDefinition
identity: iGEM#B0034
name: RBS (Elowitz 1999)
type: biopax#DnaRegion
role: SO:0000139 (Ribosome Entry Site)
ComponentDefinition
identity: iGEM#E0040
name: GFP
type: biopax#DnaRegion
role: SO:0000316 (CDS)
ComponentDefinition
identity: iGEM#B0015
name: double terminator
type: biopax#DnaRegion
role: SO:0000141 (Terminator)

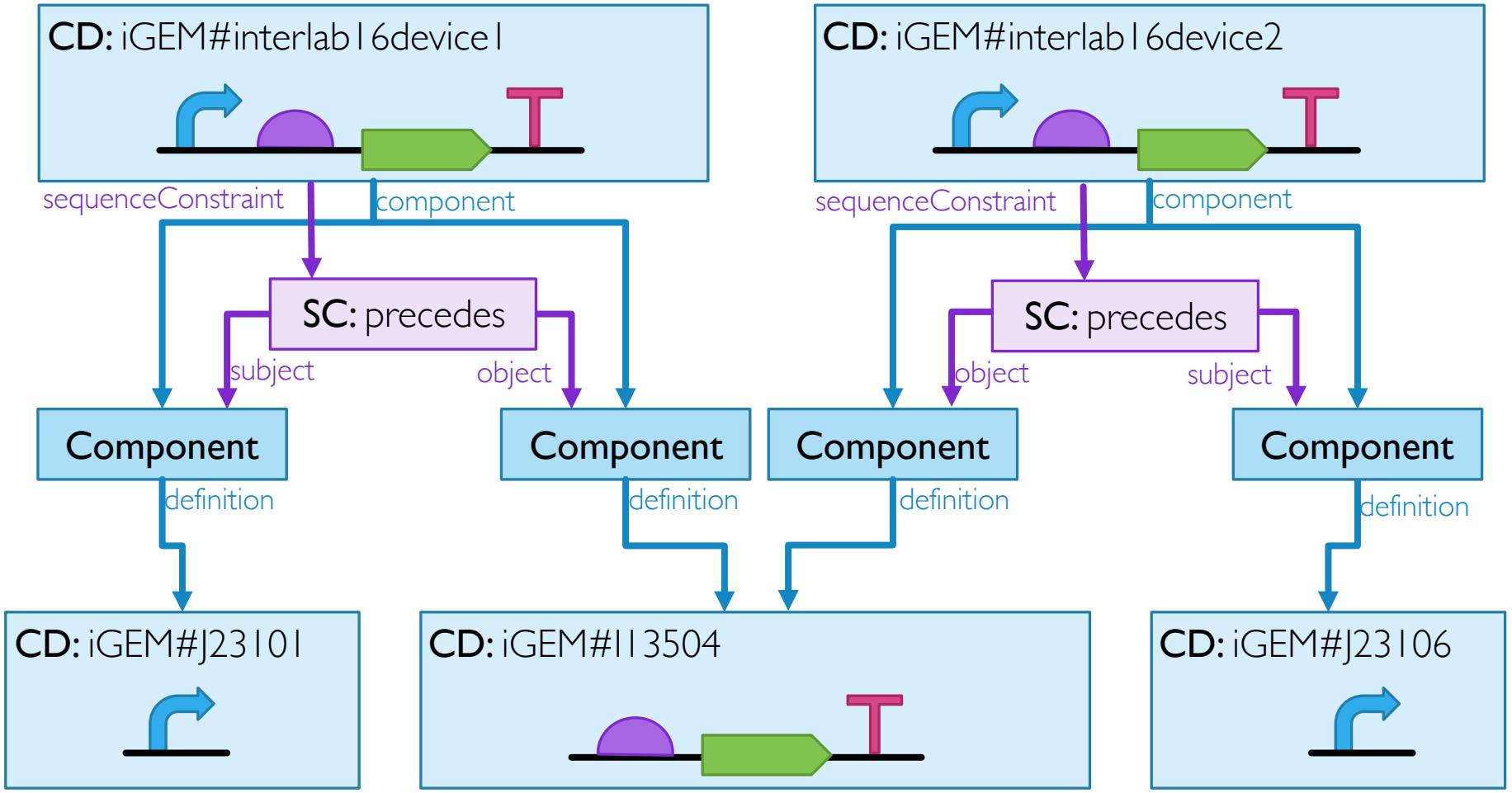
SequenceConstraint



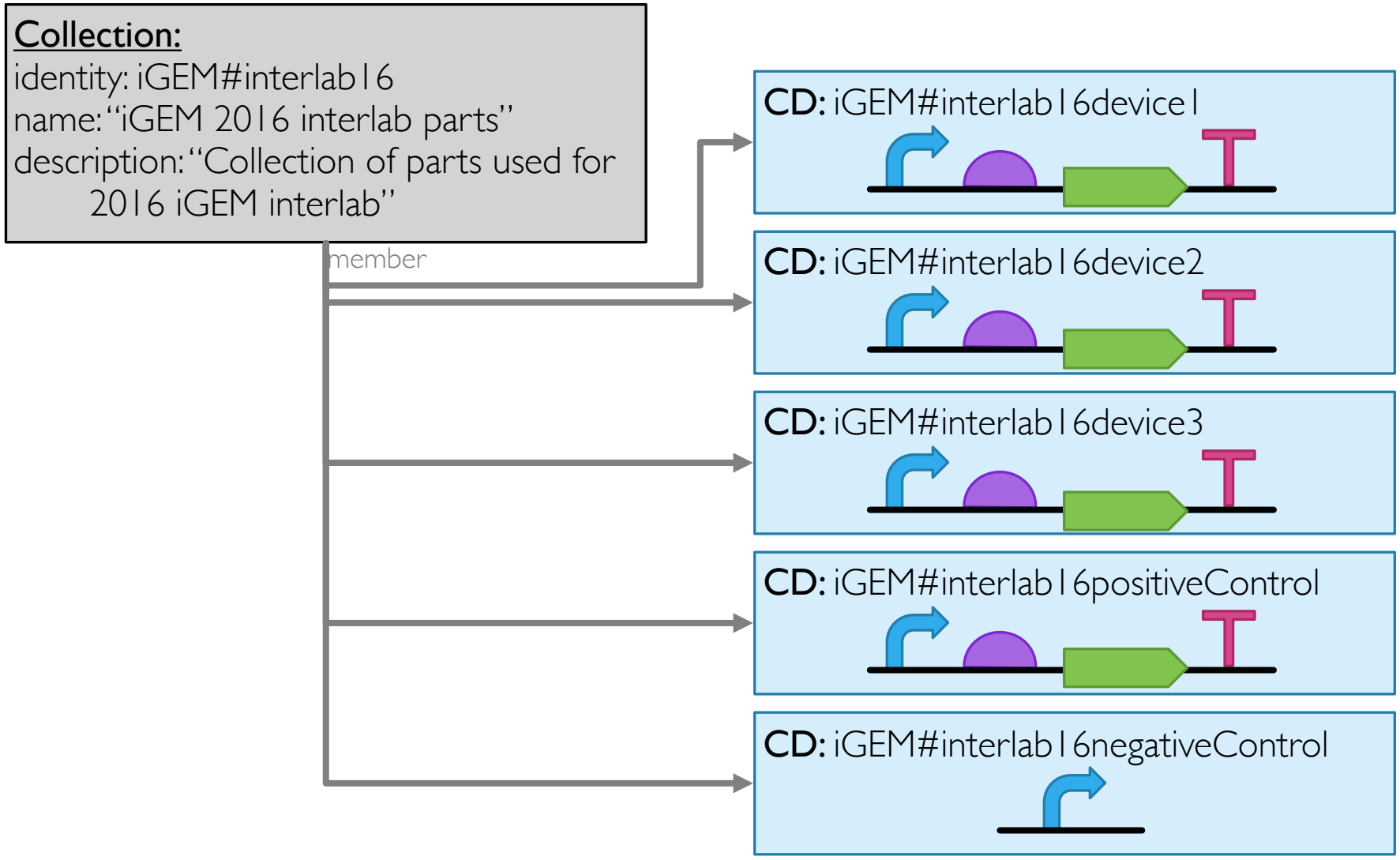
Sequence, SequenceAnnotation



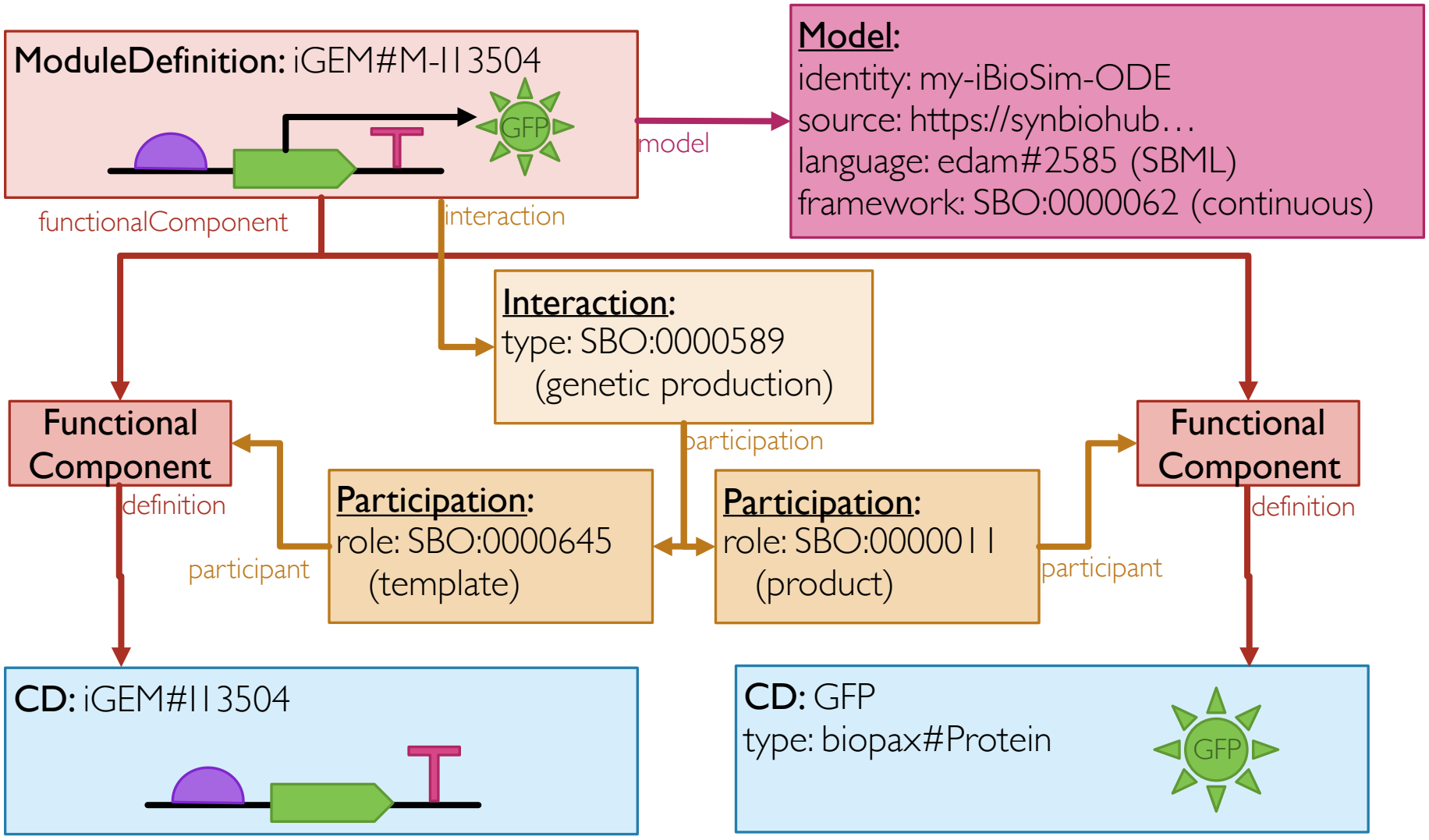
Reusing Components



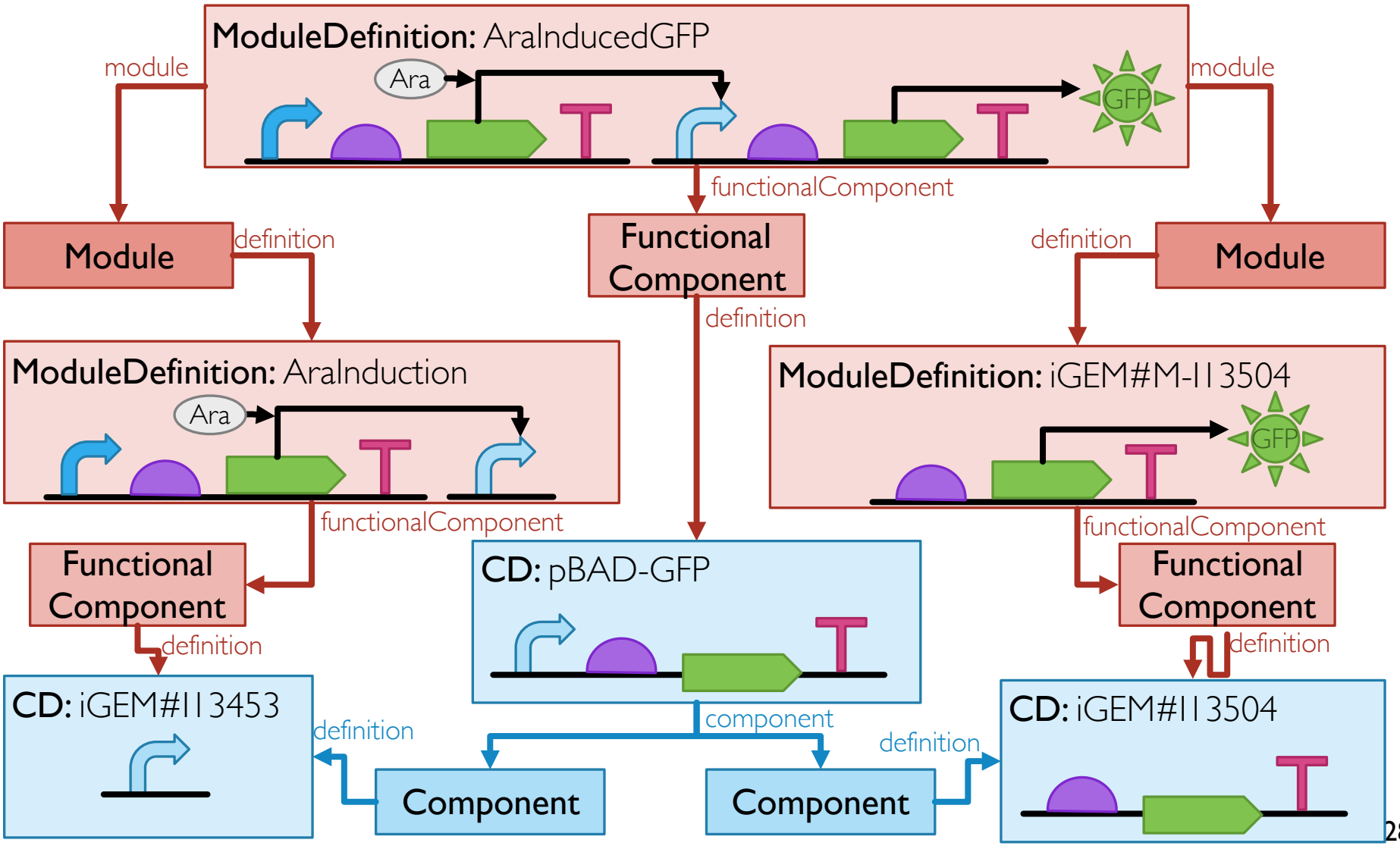
Collection


Collection:
identity: iGEM#interlab16
name: "iGEM 2016 interlab parts"
description: "Collection of parts used for 2016 iGEM interlab"
member
CD: iGEM#interlab16device1
CD: iGEM#interlab16device2
CD: iGEM#interlab16device3
CD: iGEM#interlab16positiveControl
CD: iGEM#interlab16negativeControl

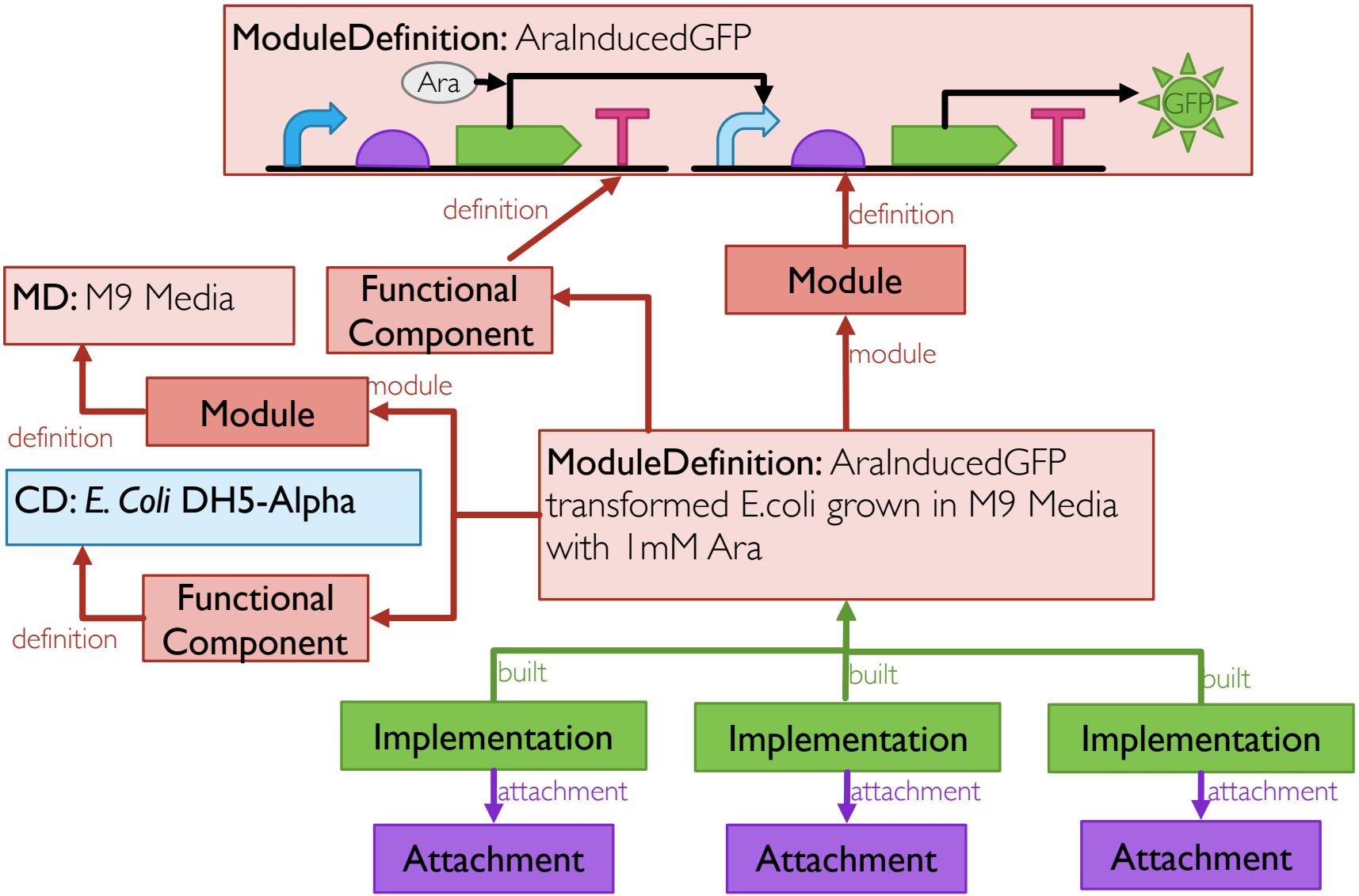
ModuleDefinition



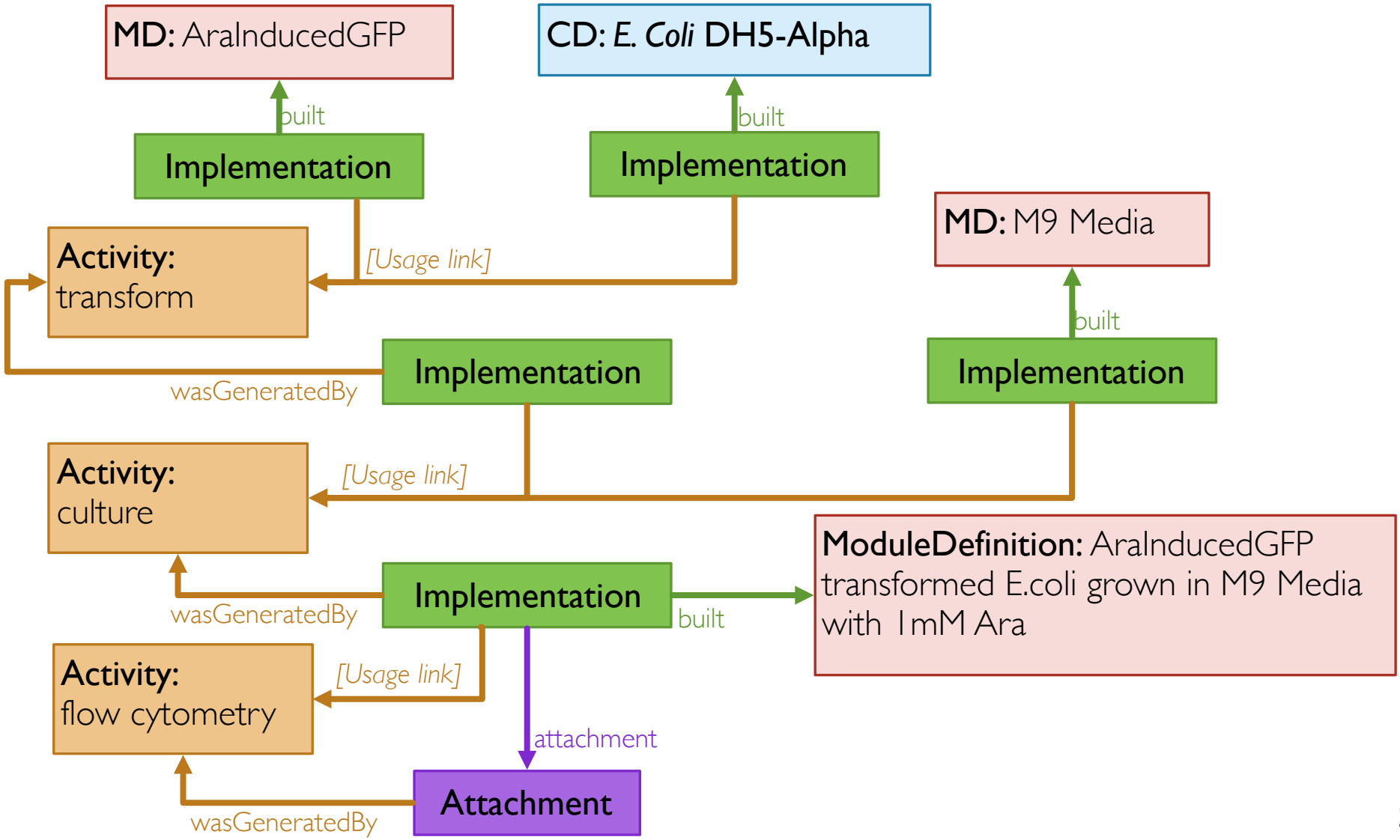
Composing Modules



Linking Designs, Protocols, and Data



Linking Designs, Protocols, and Data



Next up:

libraries & hands-on experience

Acknowledgments