# Emerging Standards in Synthetic Biology: Relationship to SBML

Herbert M Sauro

Michal Galdzicki and Deepak Chandran

University of Washington, Seattle

# What is Synthetic Biology

Synthetic Biology is the “design and construction of novel artificial biological pathways, organisms or devices, or the redesign of existing natural biological systems.” (Royal Society)

# 10 years of Synthetic Biology

Gardner TS, et al. *Construction of a genetic toggle switch in Escherichia coli.*
*Nature 2000; 403: 339-342.*

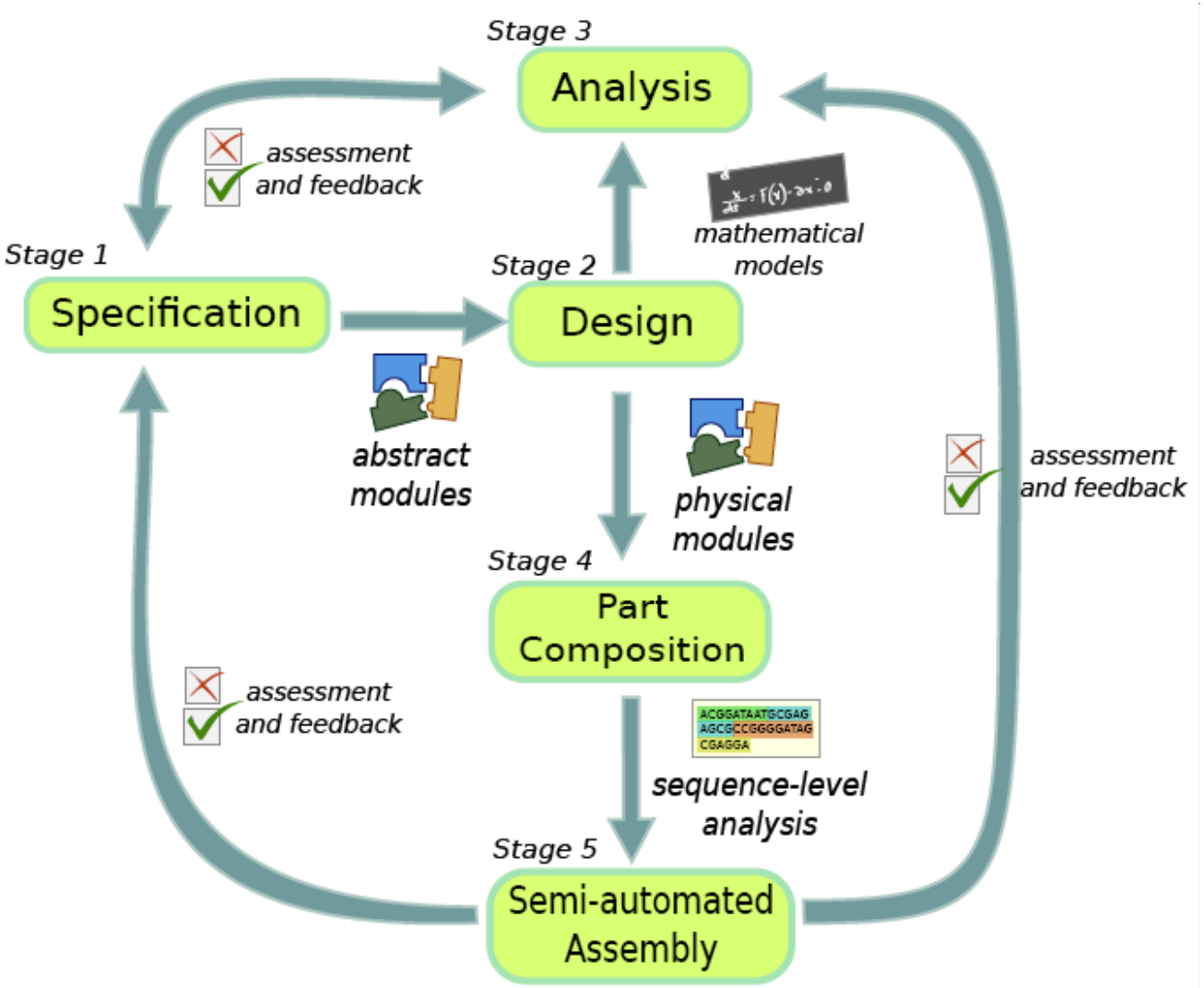
## Construction of a genetic toggle switch in *Escherichia coli*

**Timothy S. Gardner*†, Charles R. Cantor* & James J. Collins*†**

** Department of Biomedical Engineering, † Center for BioDynamics and ‡ Center for Advanced Biotechnology, Boston University, 44 Cummington Street, Boston, Massachusetts 02215, USA*

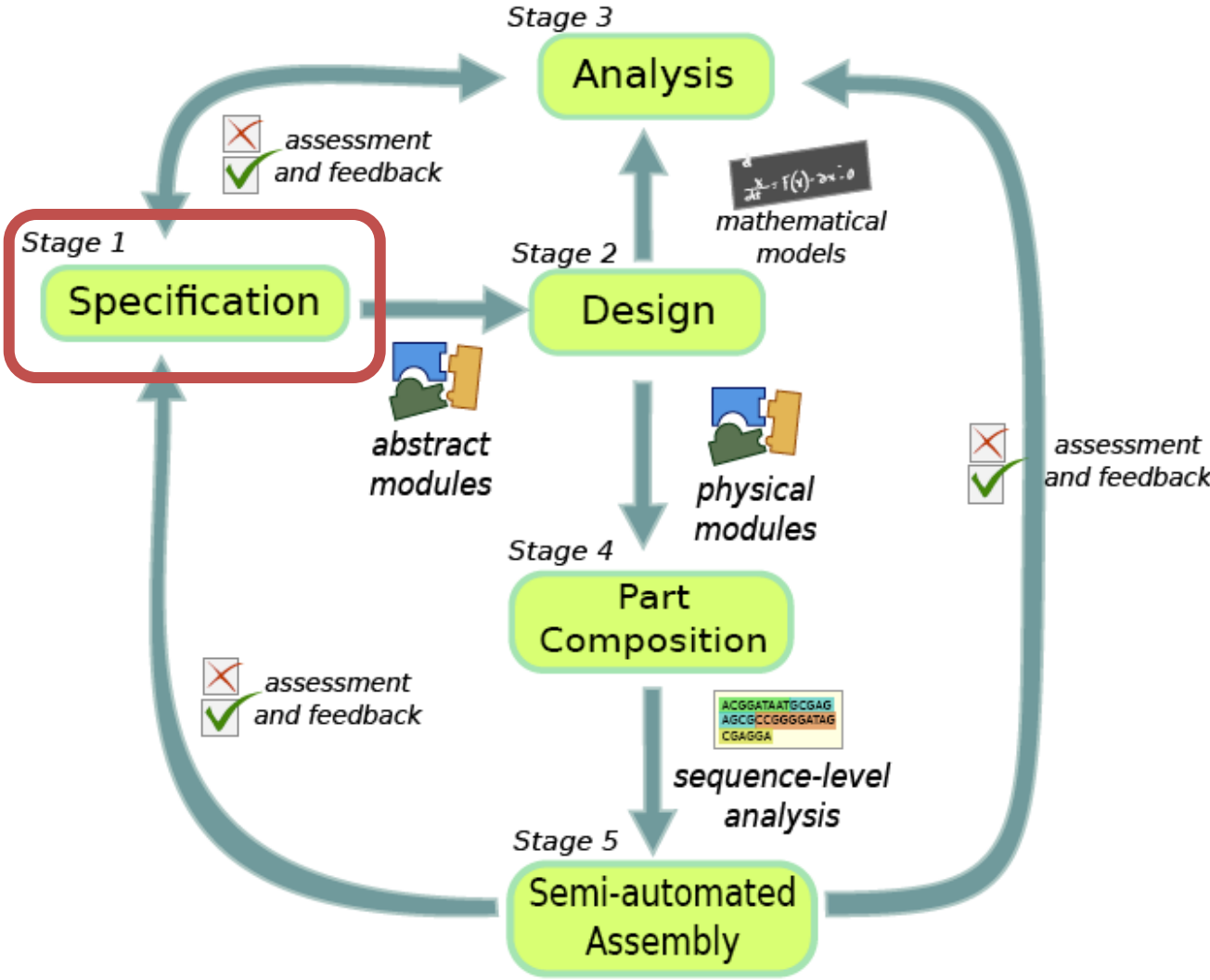
It has been proposed[1] that gene-regulatory circuits with virtually any desired property can be constructed from networks of simple regulatory elements. These properties, which include multistability and oscillations, have been found in specialized gene circuits such as the bacteriophage λ switch[2] and the Cyanobacteria circadian oscillator[3]. However, these behaviours have not been demonstrated in networks of non-specialized regulatory components. Here we present the construction of a genetic toggle switch—a synthetic, bistable gene-regulatory network—in *Escherichia coli* and provide a simple theory that predicts the conditions necessary for bistability. The toggle is constructed from any two repressible promoters arranged in a mutually inhibitory network. It is flipped between stable states using transient chemical or thermal induction and exhibits a nearly ideal switching threshold. As a practical device, the toggle switch forms a synthetic, addressable cellular memory unit and has implications for biotechnology, biocomputing and gene therapy.

# Synthetic Biology is Different from Systems Biology



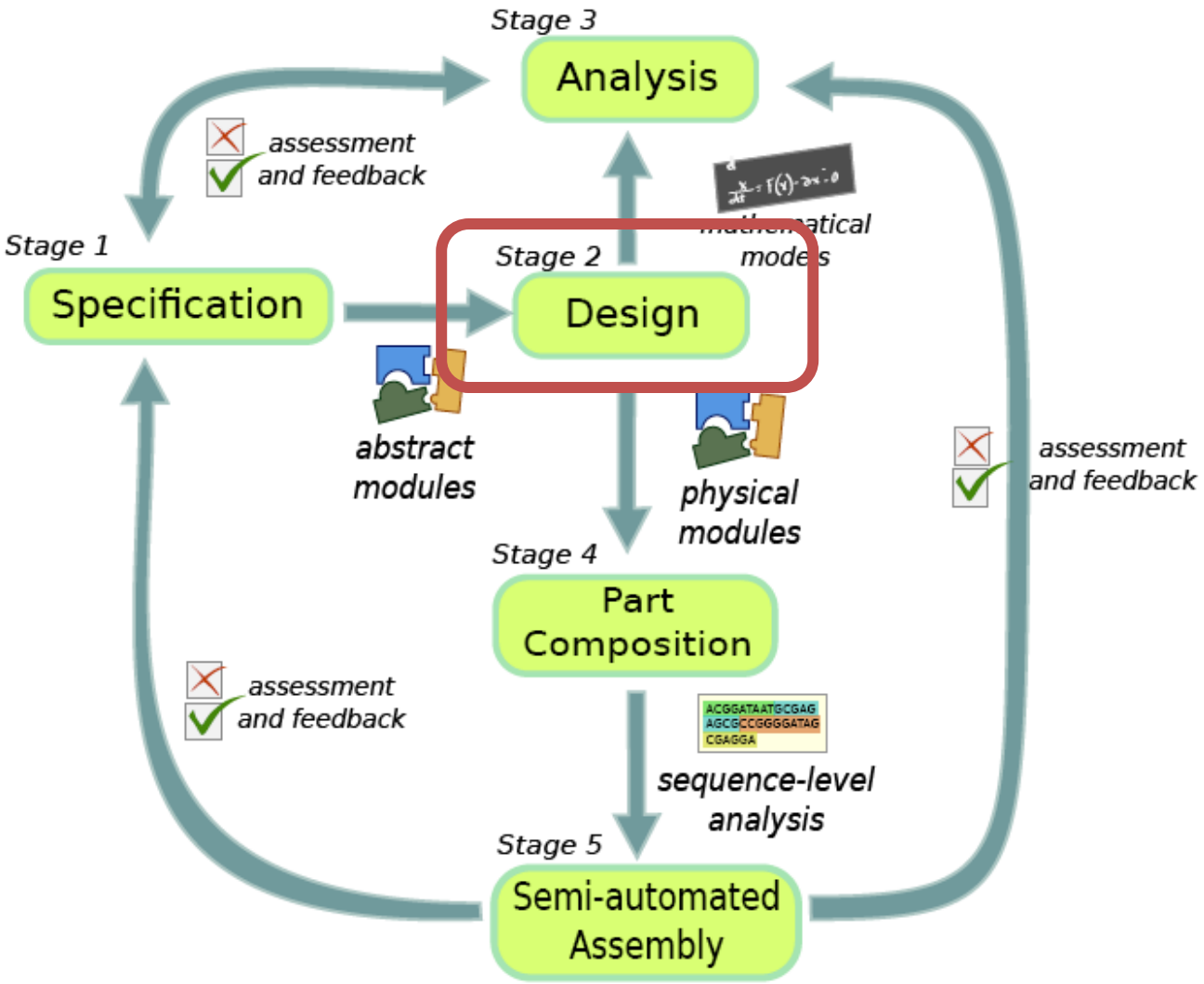
# Specification

"Creating cells that are able to excrete insulin at a specified rate during their optimal growth."



# Design

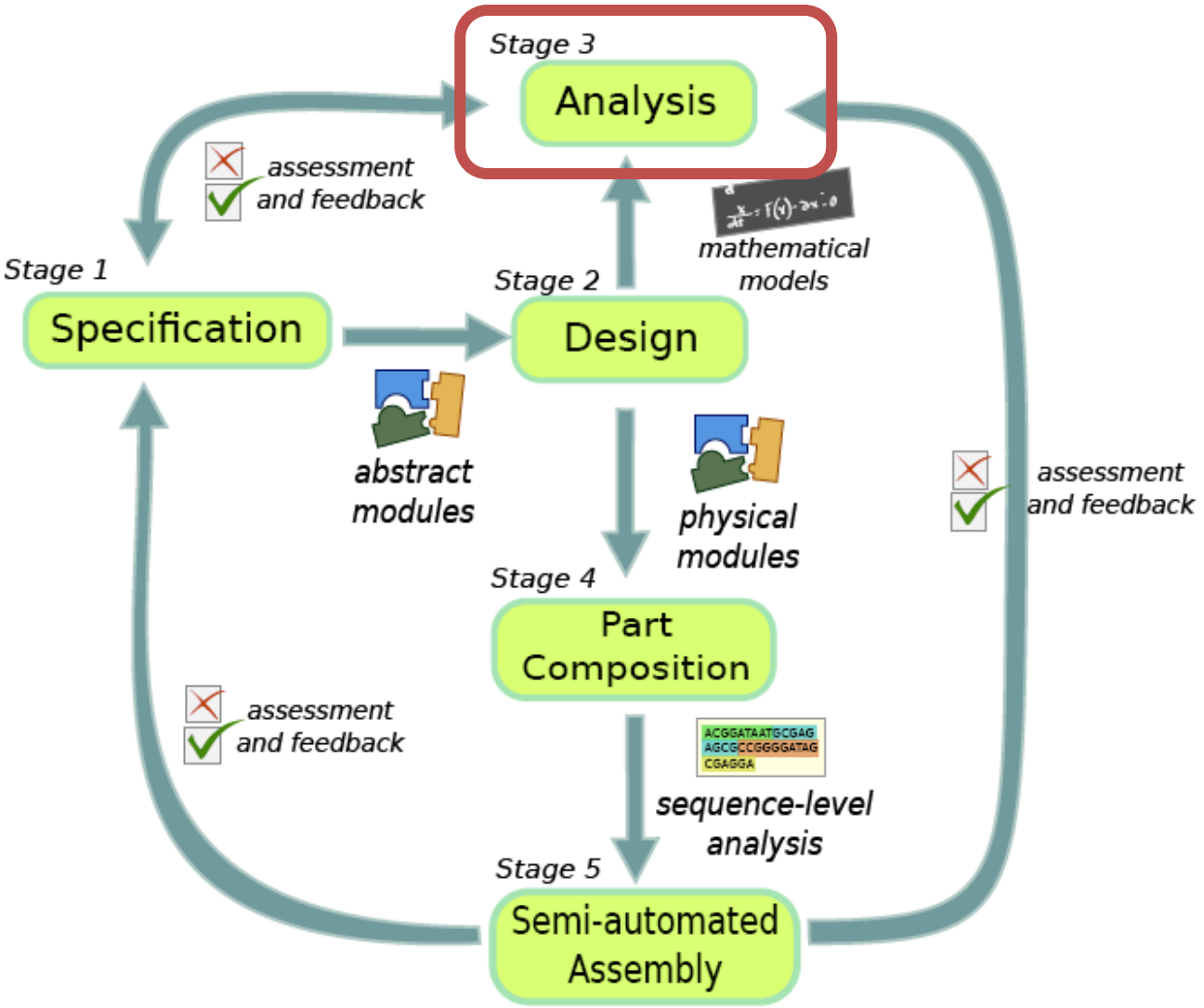
Provide the engineer with a set of building blocks that ease the task of making complex systems while ensuring that the specification is adhered to.



# Analysis

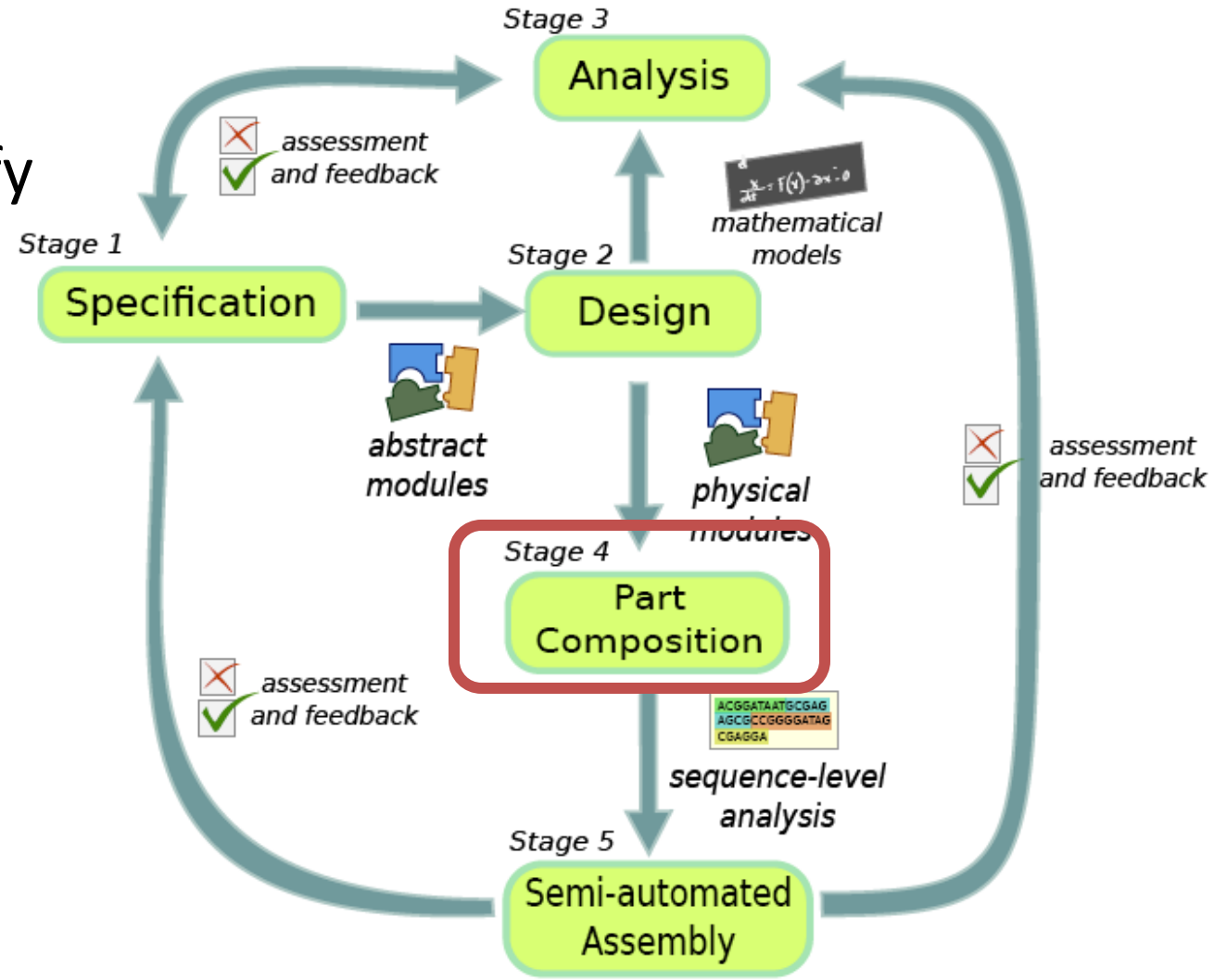
SBML/CellML

Perform analysis on a given design to quantitatively investigate how well it satisfies the specification.

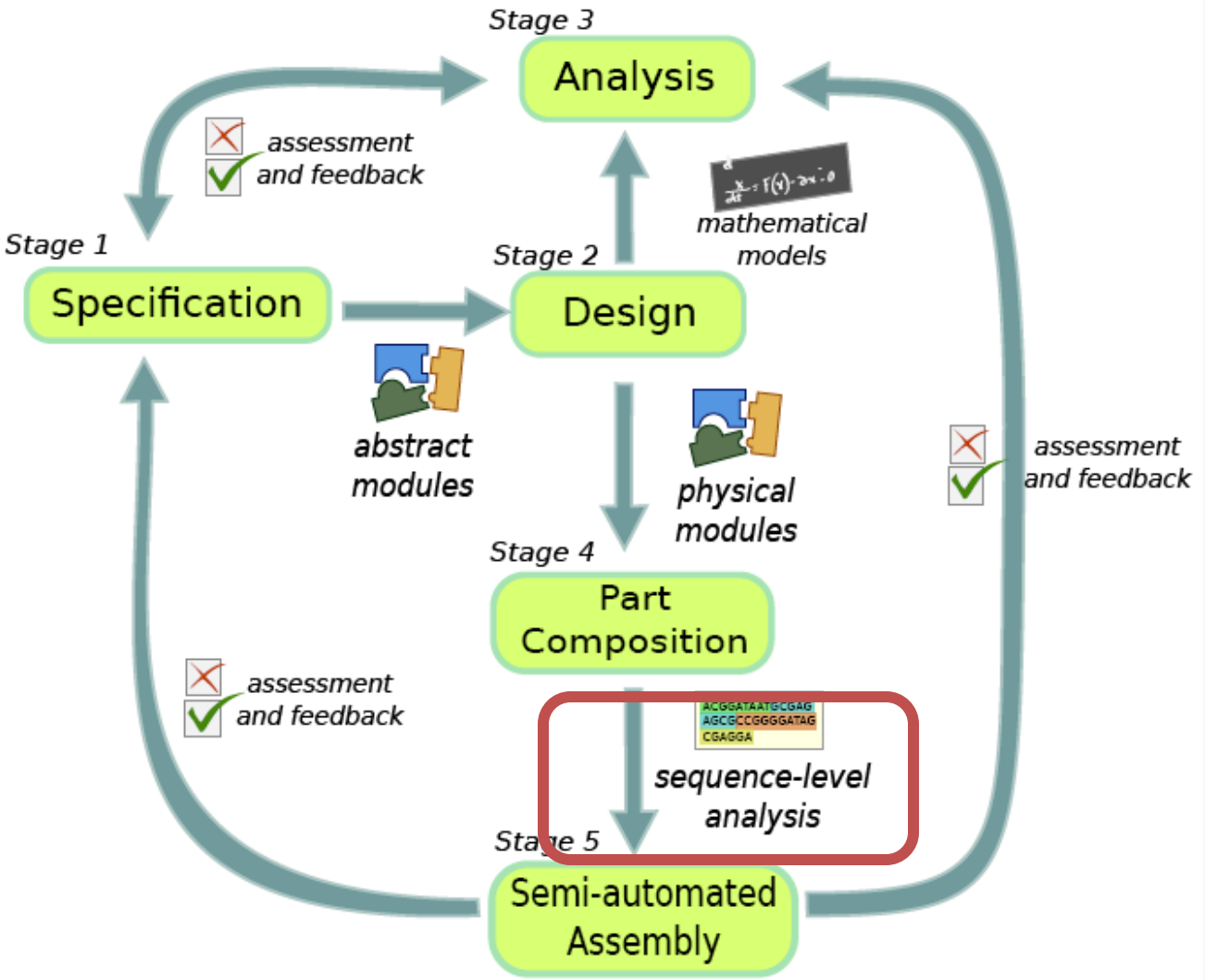


# Part Composition

Design and identify the final DNA sequence that represents the design.



# Computer Aided Design in Synthetic Biology



# Sequence Refinement

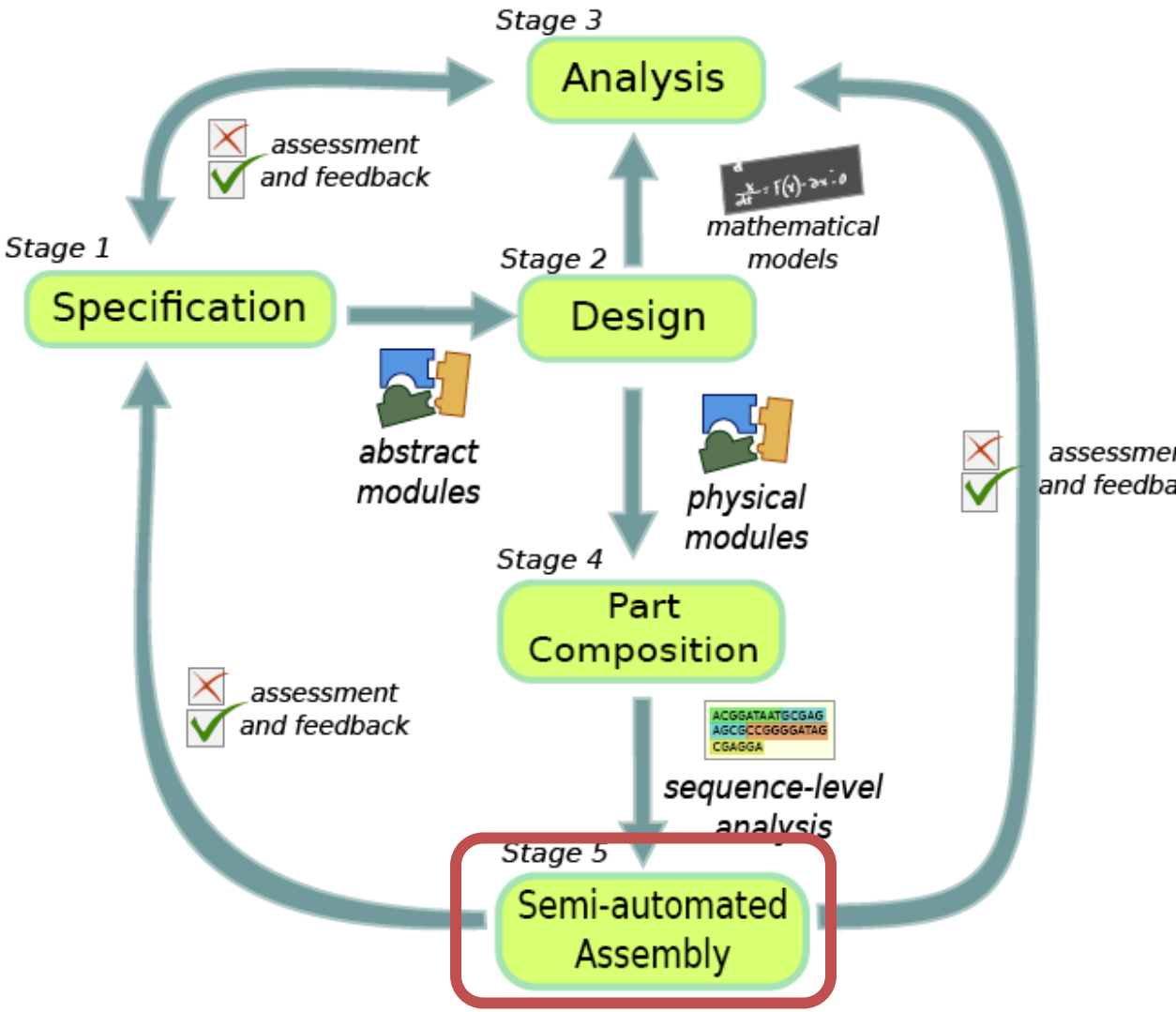
Examples include avoiding specific restriction sites or optimizing the codon usage. The sequence might also be optimized to avoid repeated regions that may be prone to recombination or secondary structures.

# Assembly

1. Biobrick standard assembly and variants: based on specific flanking sequences (Idempotent).

2. PCR based methods such as In Fusion

3. DNA synthesis



# Overall Aim of the Standardization Effort

To support the synthetic biology workflow:

1. Laboratory parts management
2. Simulation/**Analysis**
3. Design
4. Codon optimization
5. Assembly
6. Repositories - preferably distributed

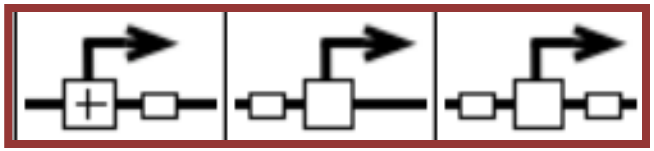
# Overall Aim of the Standardization Effort

Specifically:

- To allow researches to electronically exchange designs with round-tripping.
- To send designs to bio-fabrication centers for assembly.
- To allow storage of designs in repositories and for publication purposes.

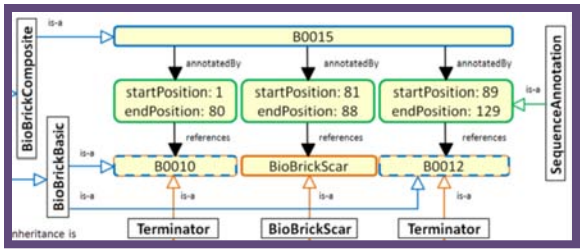
# Synthetic Biology Data Exchange Group

## Synthetic Biology Open Language (SBOL)



**SBOL-visual**

biofab




**SBOL-semantic**

UNIVERSITY of WASHINGTON

Clotho

BOSTON UNIVERSITY

Berkeley UNIVERSITY OF CALIFORNIA

TinkerCell

UNIVERSITY of WASHINGTON

GenoCAD

VIRGINIA BIOINFORMATICS INSTITUTE AT VIRGINIA TECH

jbeir public

jbei Joint BioEnergy Institute



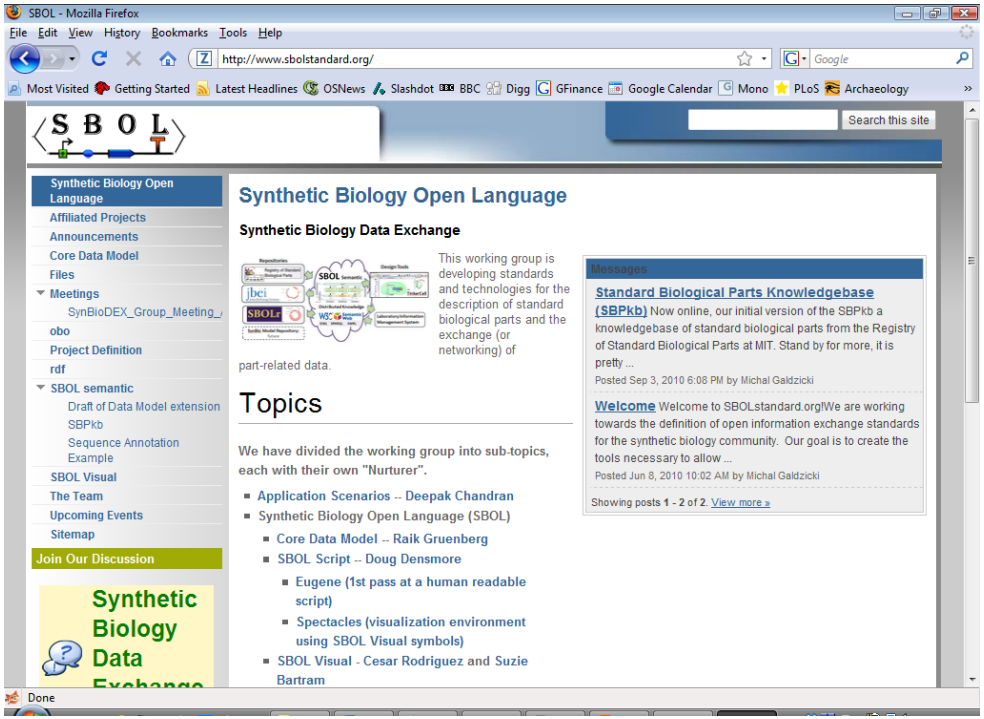
sbolstandard.org

# Synthetic Biology Open Language: SBOL

http://www.sbolstandard.org/

1. SBOL Semantic
2. SBOL Visual

Eugene Scripting from Douglas Densmore (BU)



# SBOL Visual (SBOLv)

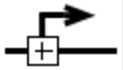
Developed at the Berkeley BIOFAB (Endy and Arkin)

- Current about 40 symbols
- Used to visually annotate ‘features’ on a DNA

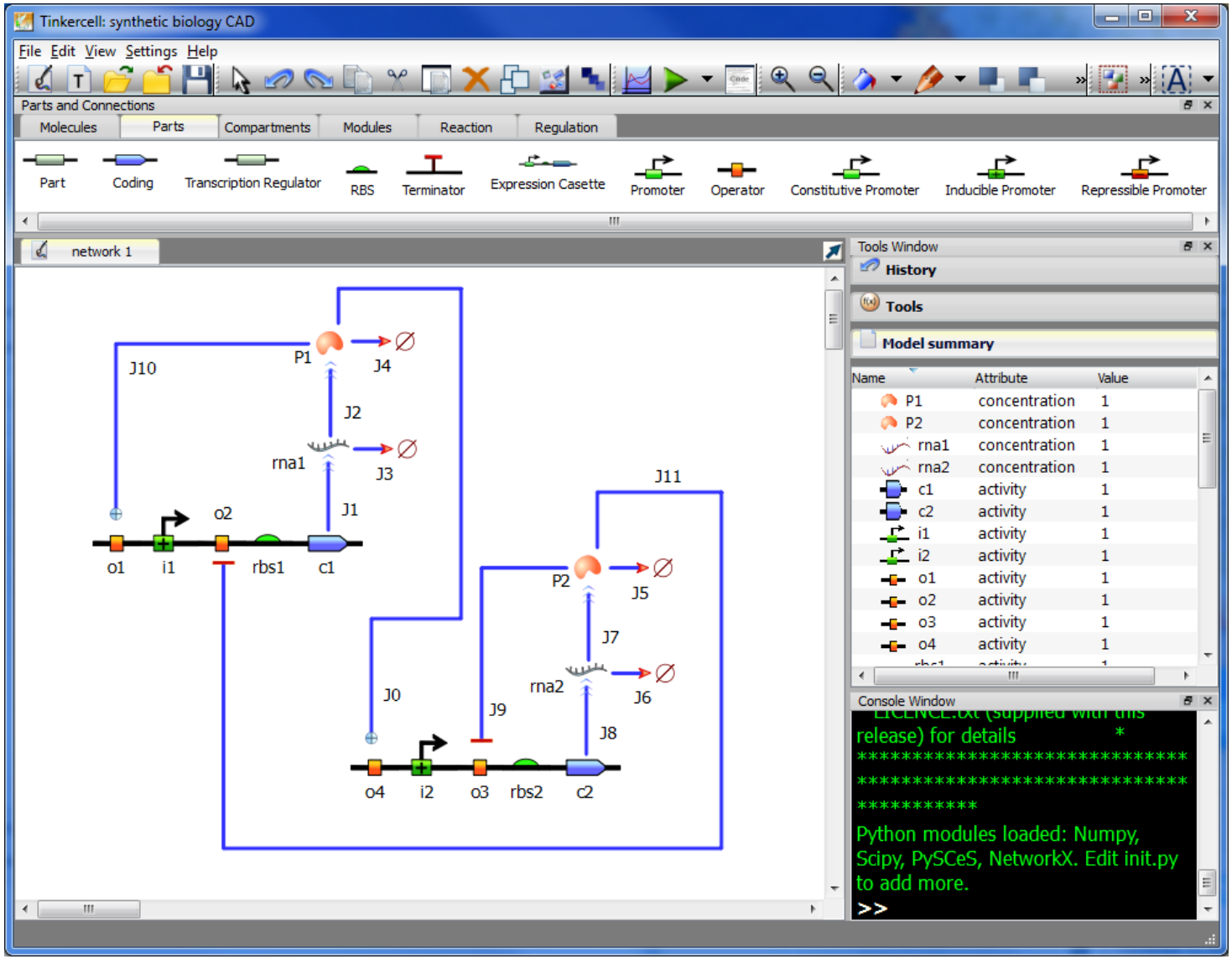
Three symbol groups at present:

1. Central Dogma
2. Genetic Engineering
3. Devices

**Central Dogma**

| Symbol | |
|---|---|
|  | **Origin of Replication** The circle represents a plasmid. The rect copy number is OPTIONAL but MUST be |
|  | **Shorthand Origin of Replication** Shorthand version of the Origin of Replicat communication. However, the Origin of Re |
|  | **Forward Constitutive Promoter** Represents a DNA sequence that prom indicates constitutive transcription. |
|  | **Reverse Constitutive Promoter** Represents a constitutive promoter initiati |
|  | **Forward Inducible Promoter** Represents a promoter that requires indu indicates that the promoter is inducible. |
|  | **Reverse Inducible Promoter** Represents a promoter that requires indu indicates that the promoter is inducible. |
|  | **Forward Repressible Promoter** Represents a promoter where transcriptio promoter is repressible. |
| | **Reverse Repressible Promoter** |

# SBOL Visual (Implemented in Spectacles and TinkerCell)



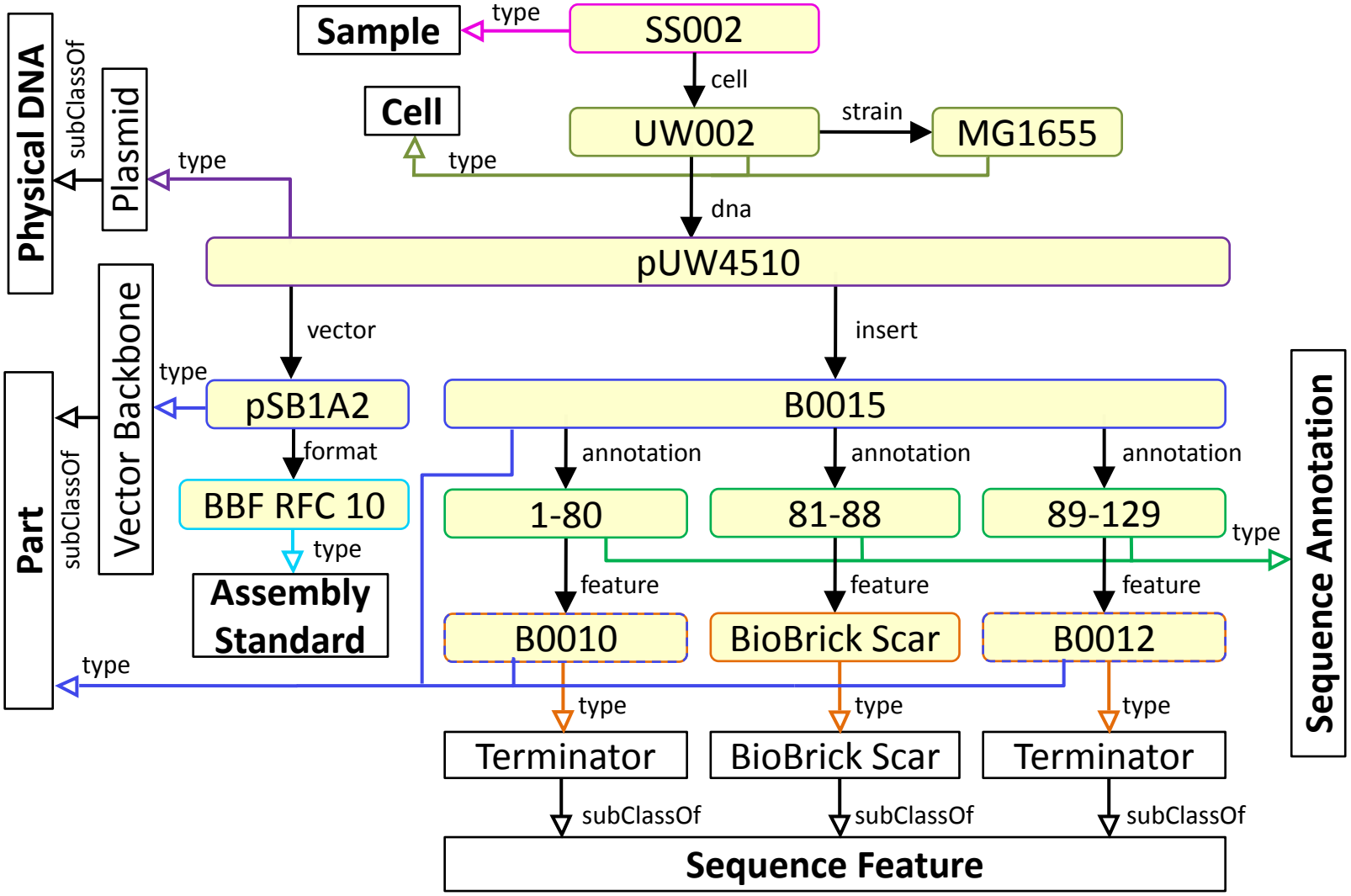



**Figure 1.** Diagram of the SBOL Semantic structure, illustrated with a set of information about a synthetic biology construct. **a.** A simplified *Class* (black rectangles) hierarchy (black open faced arrows) describes types (colored open faced arrows )of *Individual* data elements (yellow rounded rectangles) and the composition relationships between them (closed faced arrows). The example can be read as: Sample (pink) SS002 contains UW002 cells (dark green) of the MG1655 *E. coli* strain, which contain a plasmid (purple) pUW4510, which is composed of an parts (dark blue) an insert B0015 and vector backbone pSB1A2. The pSB1A2 vector backbone complies with the Assembly Standard (light blue) BBF RFC 10. The B0015 sequence annotations (green) specify three features (orange), the BioBrick Scar, and the parts (blue and orange indicating multiple inheritance), B0010 and B0012, which serve as transcriptional termination signals. **b.** Data type *Properties* used to hold information for each SBOL class follow the colon.

# Eugene Script

```
// Example
// Douglas Densmore

Property sequence(txt);
Property name(txt);
Property relativeStrength(num);
Property Neg10Neg35(txt[]);

Part Promoter(name, sequence, relativeStrength);
Part SpecialPromoter(name, sequence, Neg10Neg35);

Promoter p1("PromoterType1", "ATC", 10);
Promoter p2(.sequence("ATCCCCC"));
SpecialPromoter p3 ("PromoterType2", "CAT", ["CAT", "TAG"]);

Device d1(p2, p3, p1);

print(p1.sequence);
print(p3.name);
print(d1[0].sequence);
```

# List of Software

- **Clotho (BU,Berkeley), Java**

  Connects users to repositories of biological parts. Plugin tools then define the various functions that can be performed, mainly related to lab and parts managements.

- **GenoCAD (VT)**

  Web based tool for the design of biological devices using an attribute grammar which defines the legal composition of parts.

- **JBEI Repository (Java), DOE support infrastructure**

- **SBOLr (Currently Python but moving to Java)**

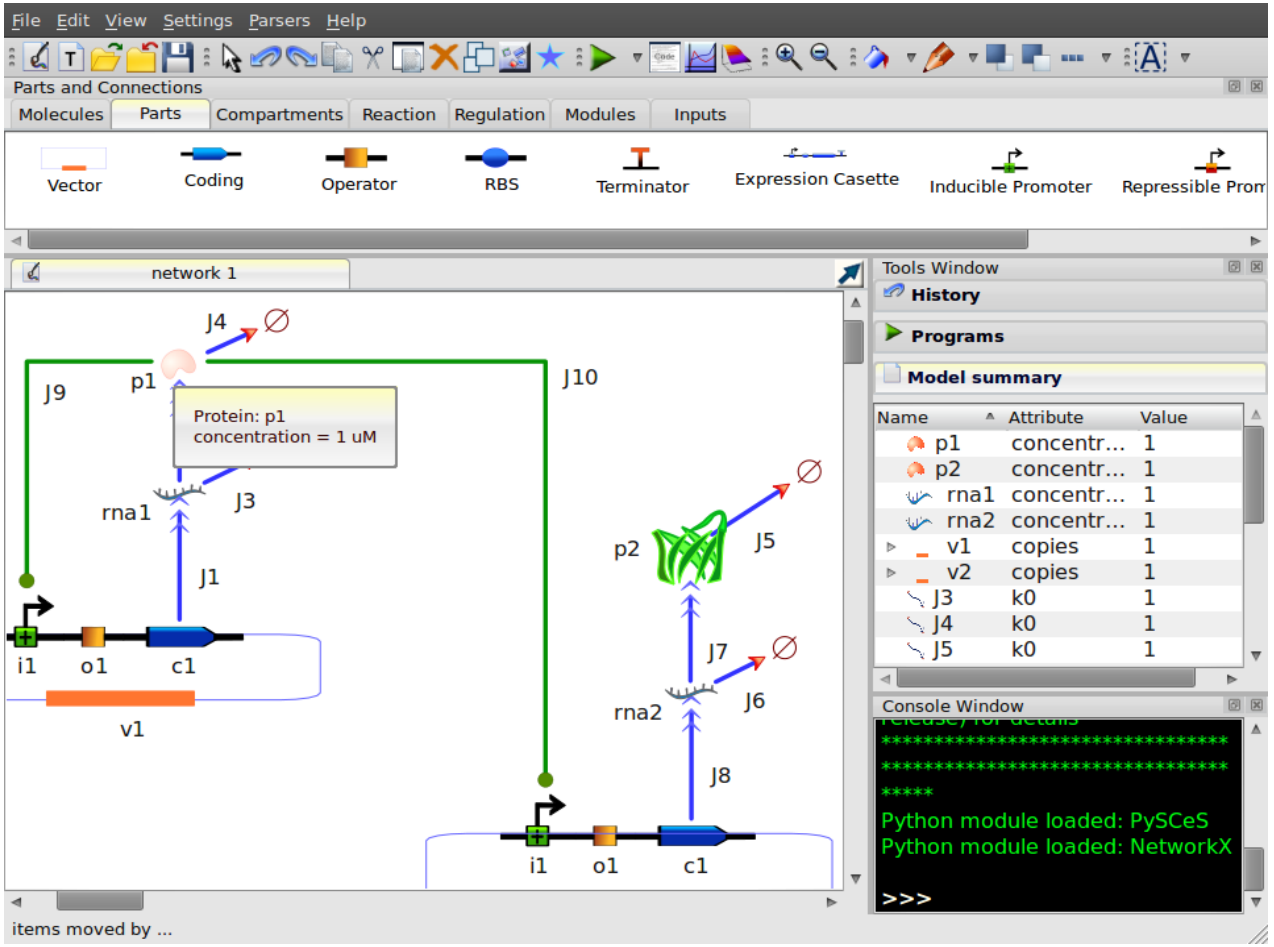
  Test parts repository, ~5000 parts

- **TinkerCell (UW), C++/Qt**

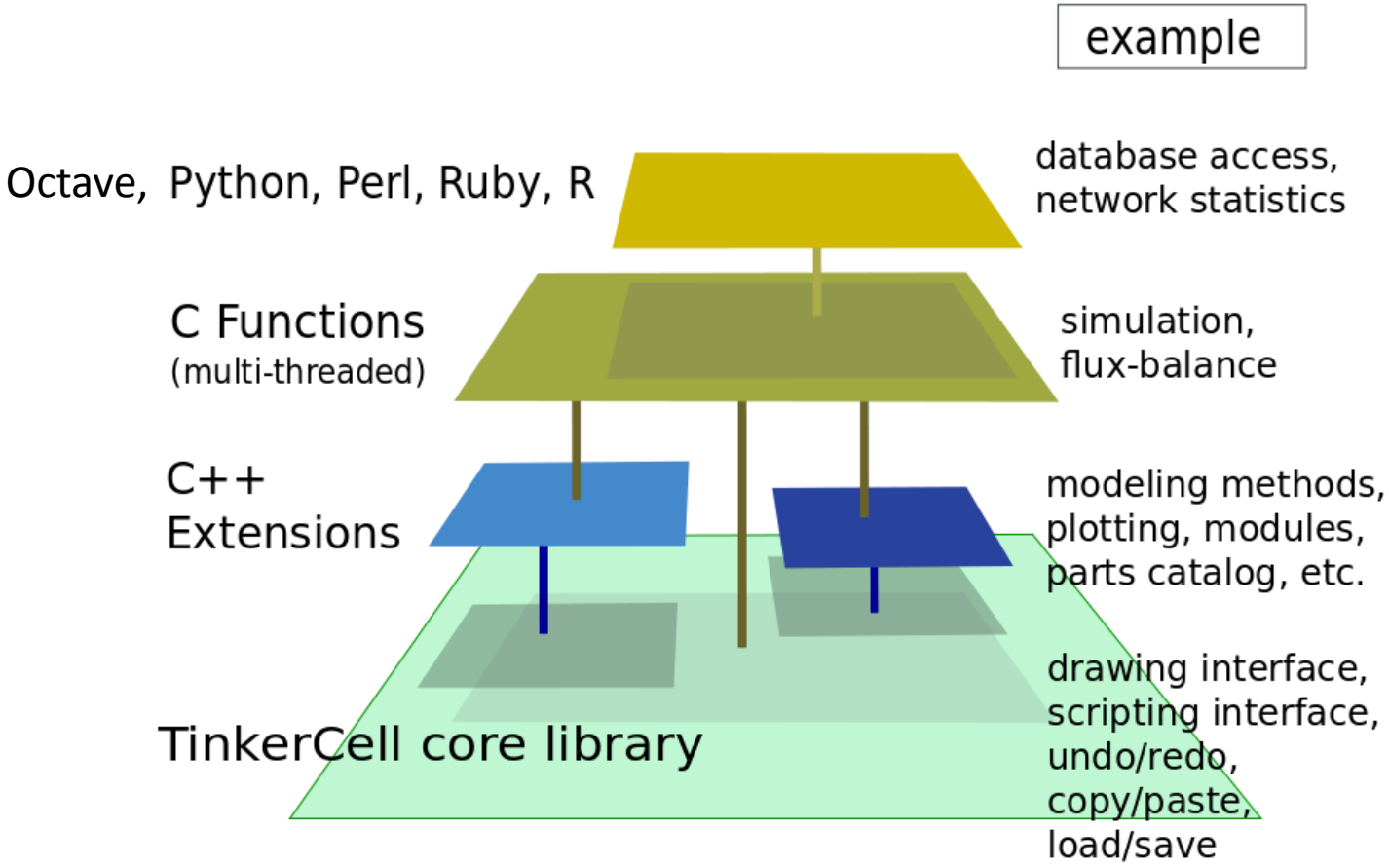
  Extensible visual design tool with support for modules, eg simulation, annotation.

# TinkerCell: Project to explore the potential of computer aided design in synthetic biology

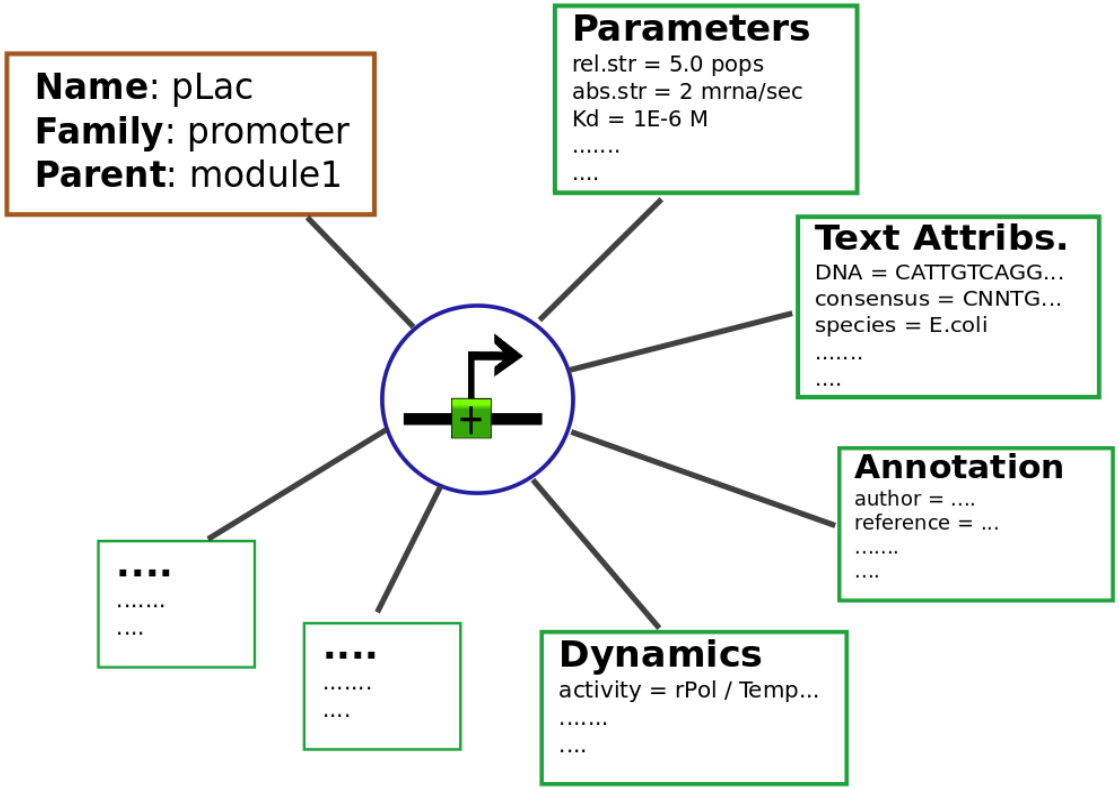
First prototype called Athena developed by Bergmann and Chandran



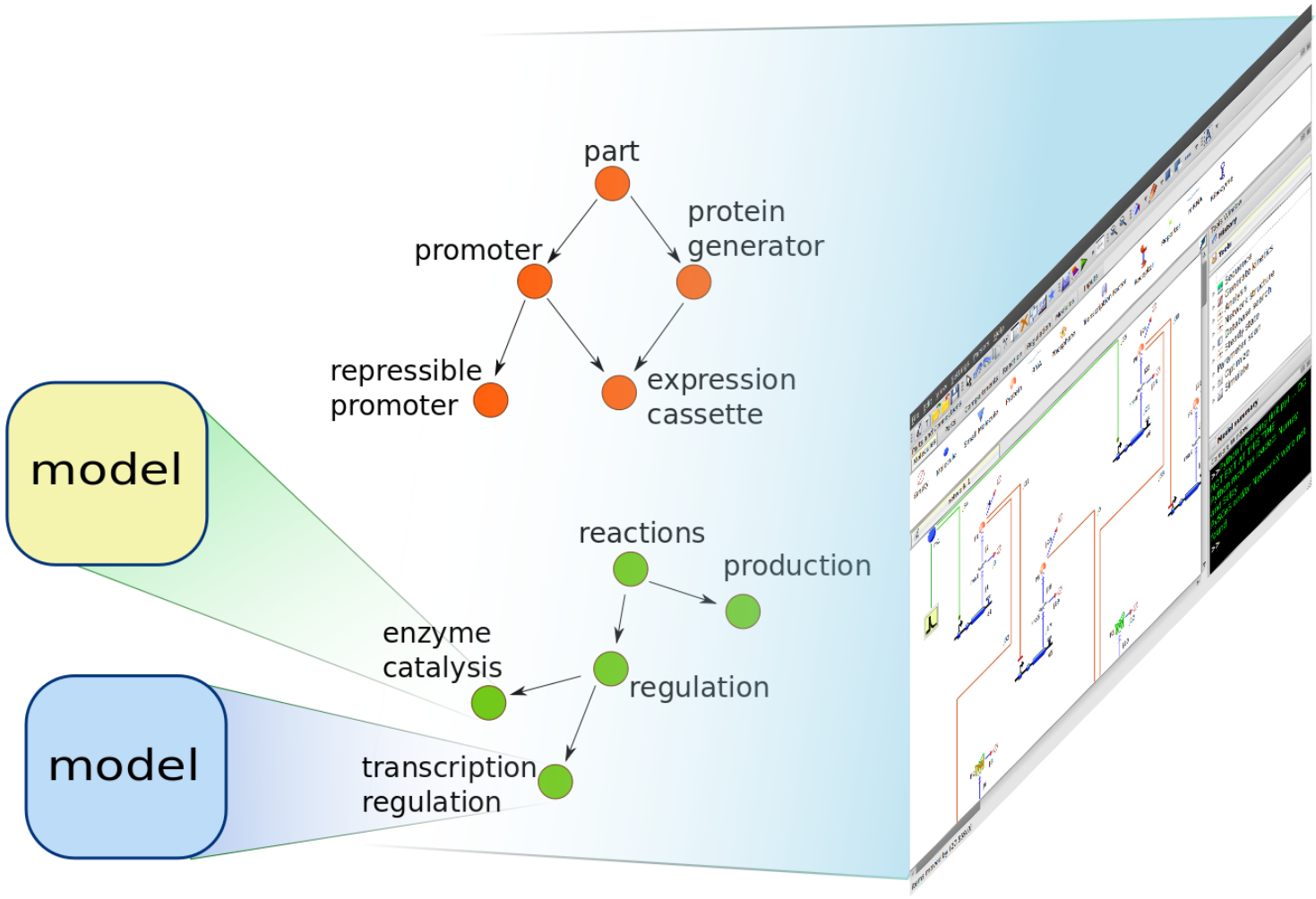
# Layered Architecture: Based on C++/Qt



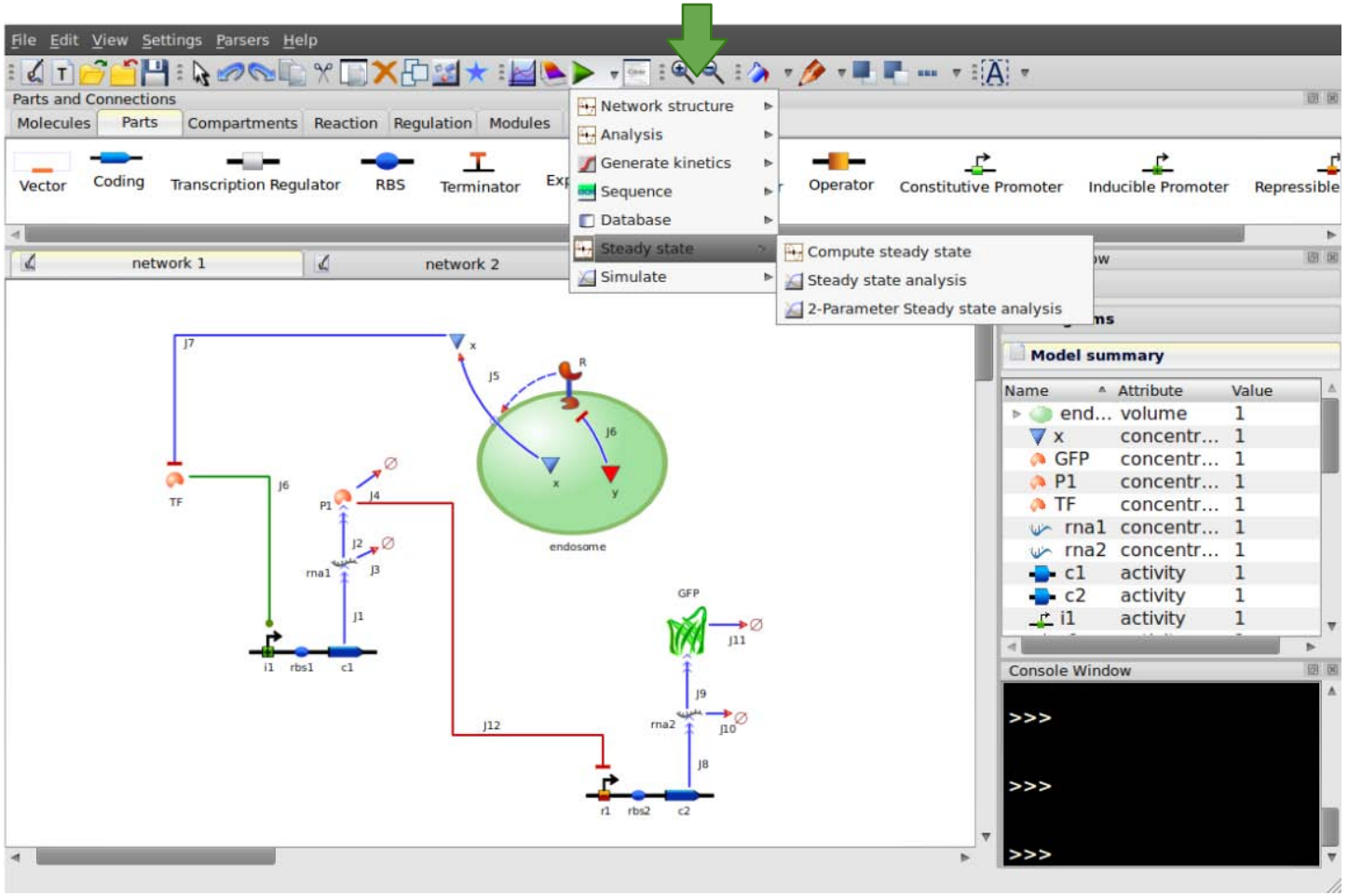
# Each component in the TinkerCell diagram is associated with one or more tables



# TinkerCell supports hierarchical sub-models

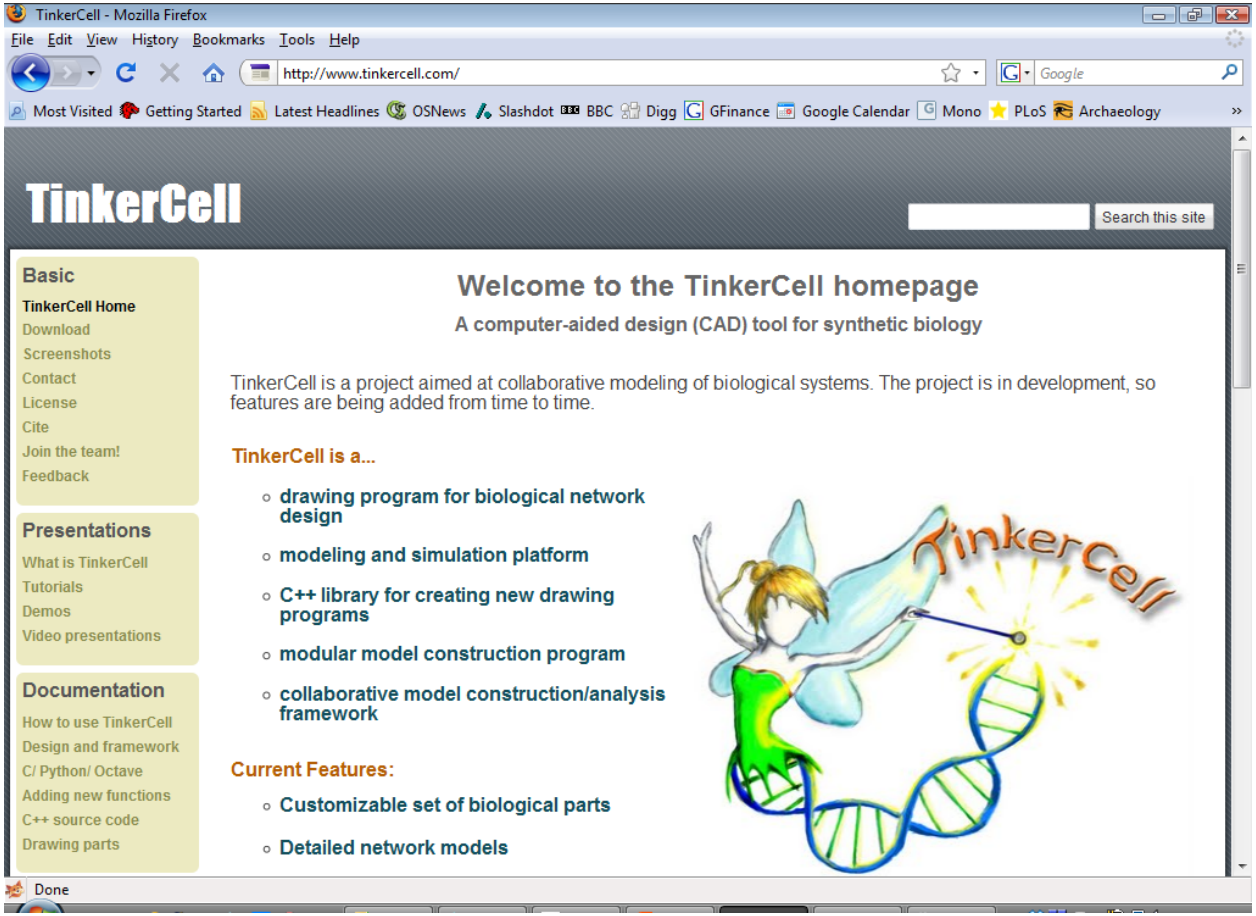


# Plugins can be written using C++, C, Octave, or Python



# Availability

www.tinkercell.com (Windows, Mac and Linux, released under BSD)
Contact author for details (dchandran1@gmail.com)



# C and Python packages included with TinkerCell

## Sundials

- Time-course simulation
- Steady state plot

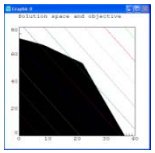
## PySCeS

Python Simulator for Cellular Systems

- structural analysis
- sensitivity analysis
- bifurcation analysis
- parameter scan
- simulation

## Custom programs

- Gillespie algorithm
- Hill equation derivation
- 2^N automatic binding events
- Loops in Jacobian



## lp_solve

- flux balance analysis

## SciPy

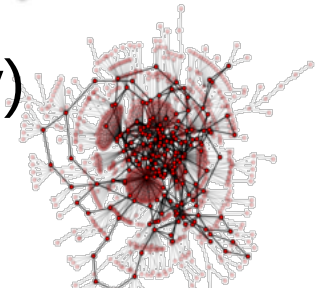
- optimization
- matrix operations
- statistics
- numerical methods
- frequency analysis

## Antimony

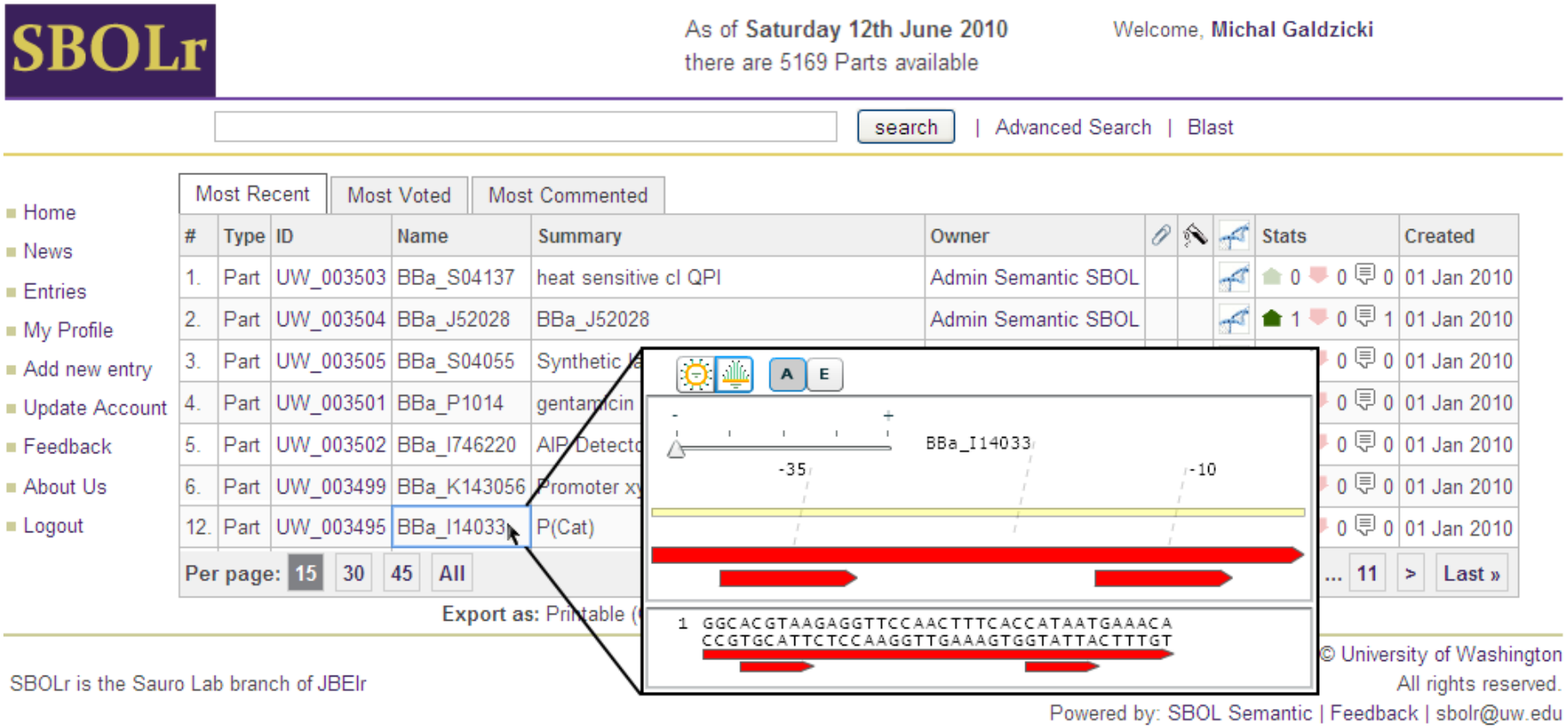
SBML and CellML support

## NetworkX (lanl.gov)

- graph analysis
- graph layout



# SBOLr is a web front end for a knowledge base of biological parts




http://sbolr.bhi.washington.edu/

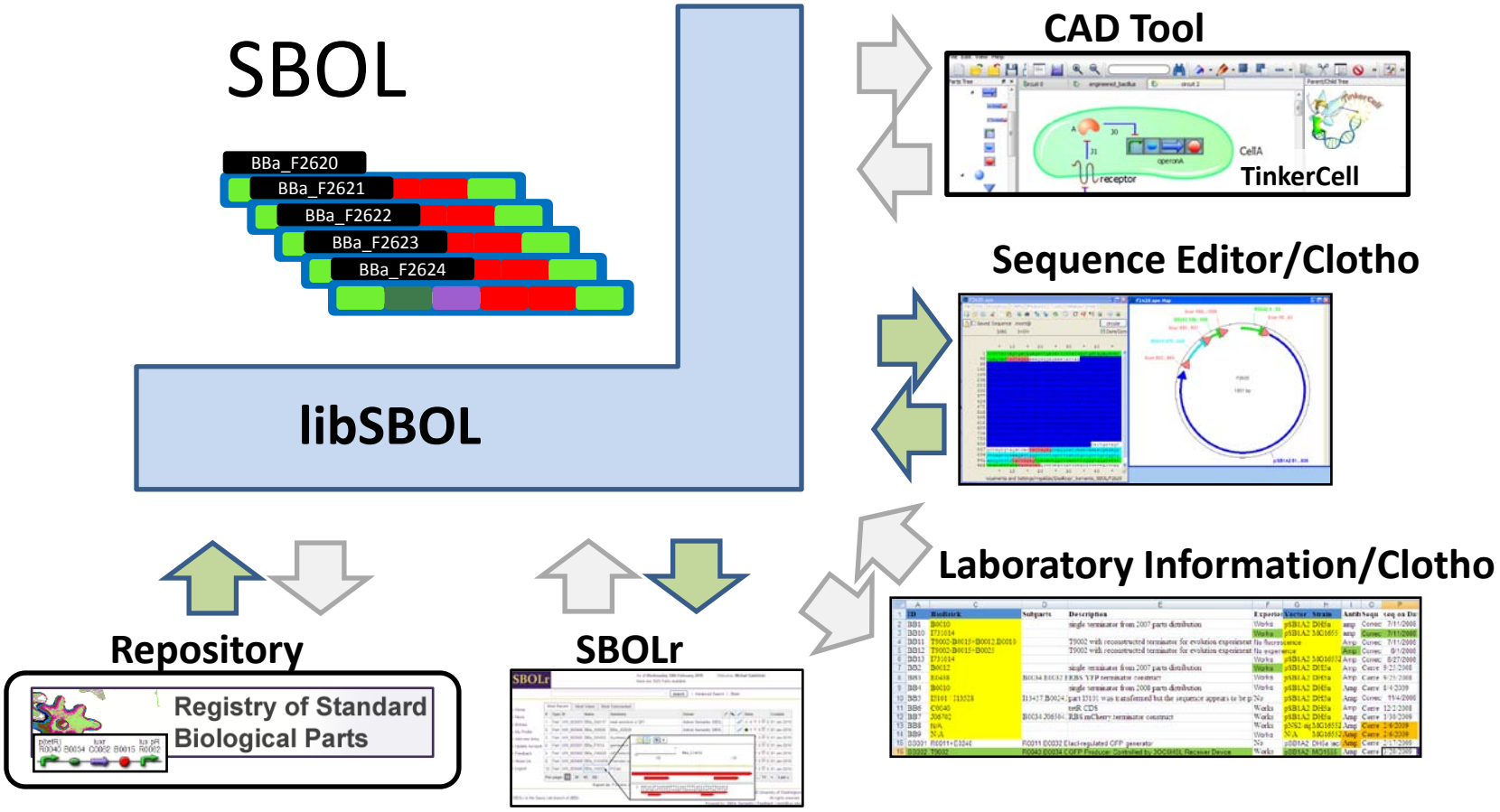


JBEIr Software by Timothy Ham & Zinovii Dmytriv

# Demo

Building: 0.00
Composition: 1.48

# libSBOL: Supports SBOL format for other third-party tools



# Relationship to Efforts in the Systems Biology Community

Standards that may have utility in the synthetic biology field:

- SBML
- CellML
- SBGN
- SEDML
- BioPAX

# SBML

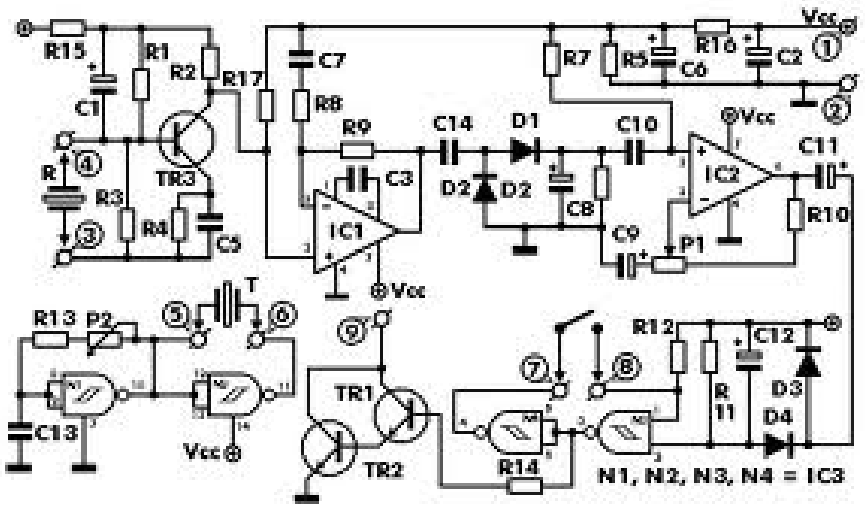
SBOL Semantic could be made into a SBML package, but no resource to currently undertake this task.

More importantly SBOL Semantic needs to be exercised more before it can be formalized into SBML. Eg Successful round tripping must be established first.

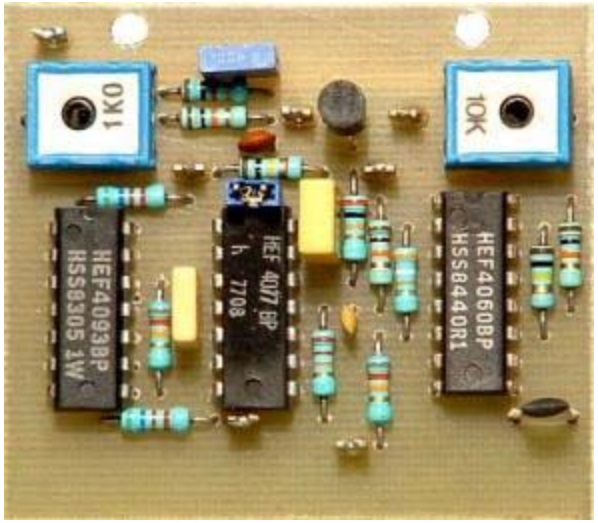
# SBGN

In some respects SBGN is too detailed but on the other hand lacks sufficient support for genetic networks. What would be desirable would be a SBGN-lite version that gives a higher level view but with gene regulatory support. Ideally, users should be able to drill down to view more or less detail.

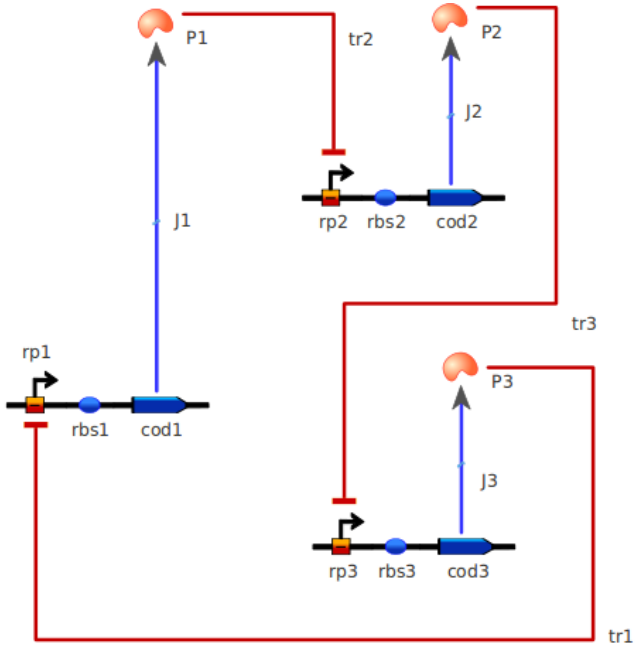
We need this:



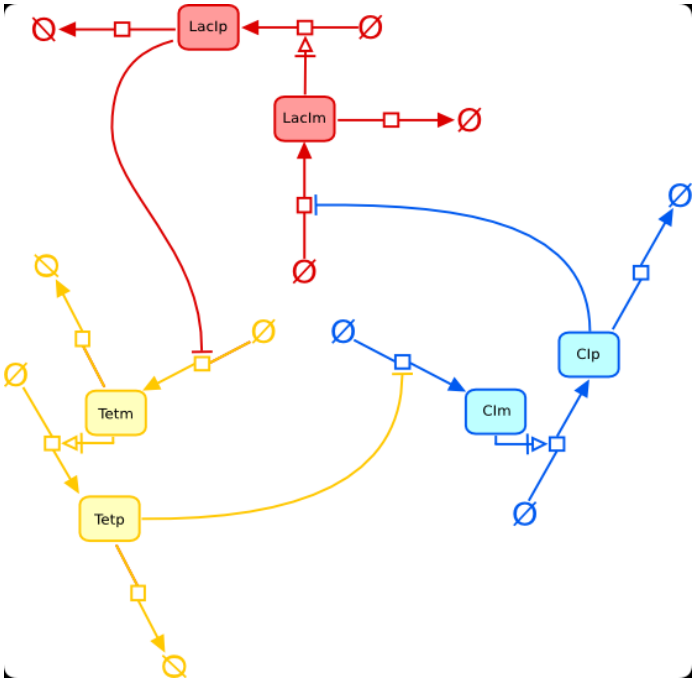

But SBGN is more like this:



# TinkerCell/SBOLv




# SBGN



# Acknowledgements: Collaborators and Funding


**UW**

Deepak Chandran, Michal Galdziki, Sean Sleight, Bryan Bartley, Alex Neilson

**BioFAB (Berkeley, Stanford)**

Cesar Rodriguez, Drew Endy , Chris Anderson

**Virginia Tech**

Jean Peccoud

**JBEI (DOE)**

Timothy Ham & Zinovii Dmytriv

**Boston University**

Douglas Densmore

**CRG, Spain**

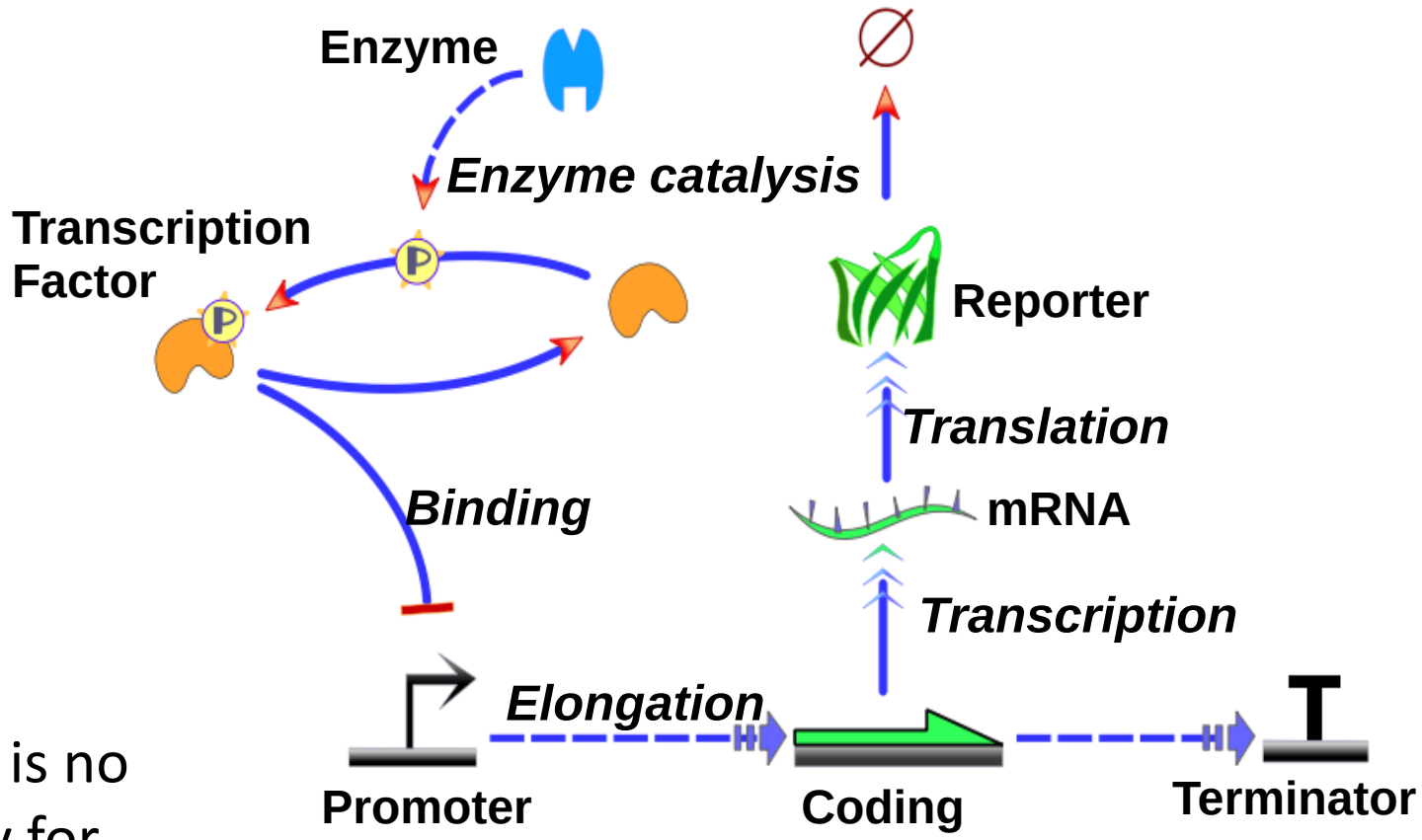
Raik Grunberg

**Funding**

National Library of Medicine (Galdziki)

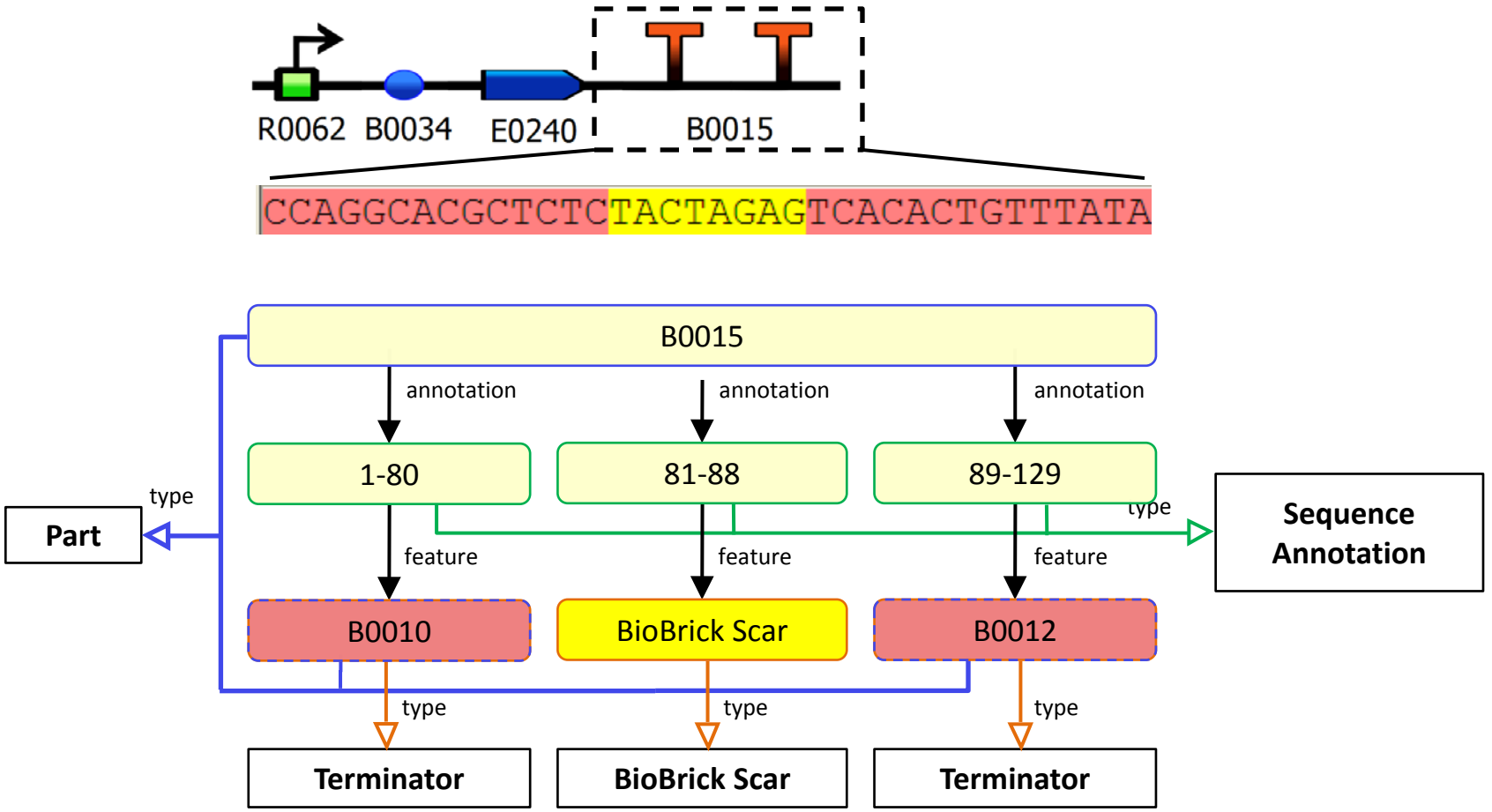
Other funding scraped from NIH (JSIm/SBW) , NIH (SBW) and NSF (FIBR), Microsoft Strategic Research, BioFAB, JBEI

# The Table Structure Allows Parts to be Semantically Annotated. This allow plug-ins to interpret a given design.




Currently there is no formal ontology for synthetic biology but one will need to be developed.

# The SBOL semantic structure




type = instance-of

Transmitted in the form of RDF/XML, using OWL schema