

# 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***

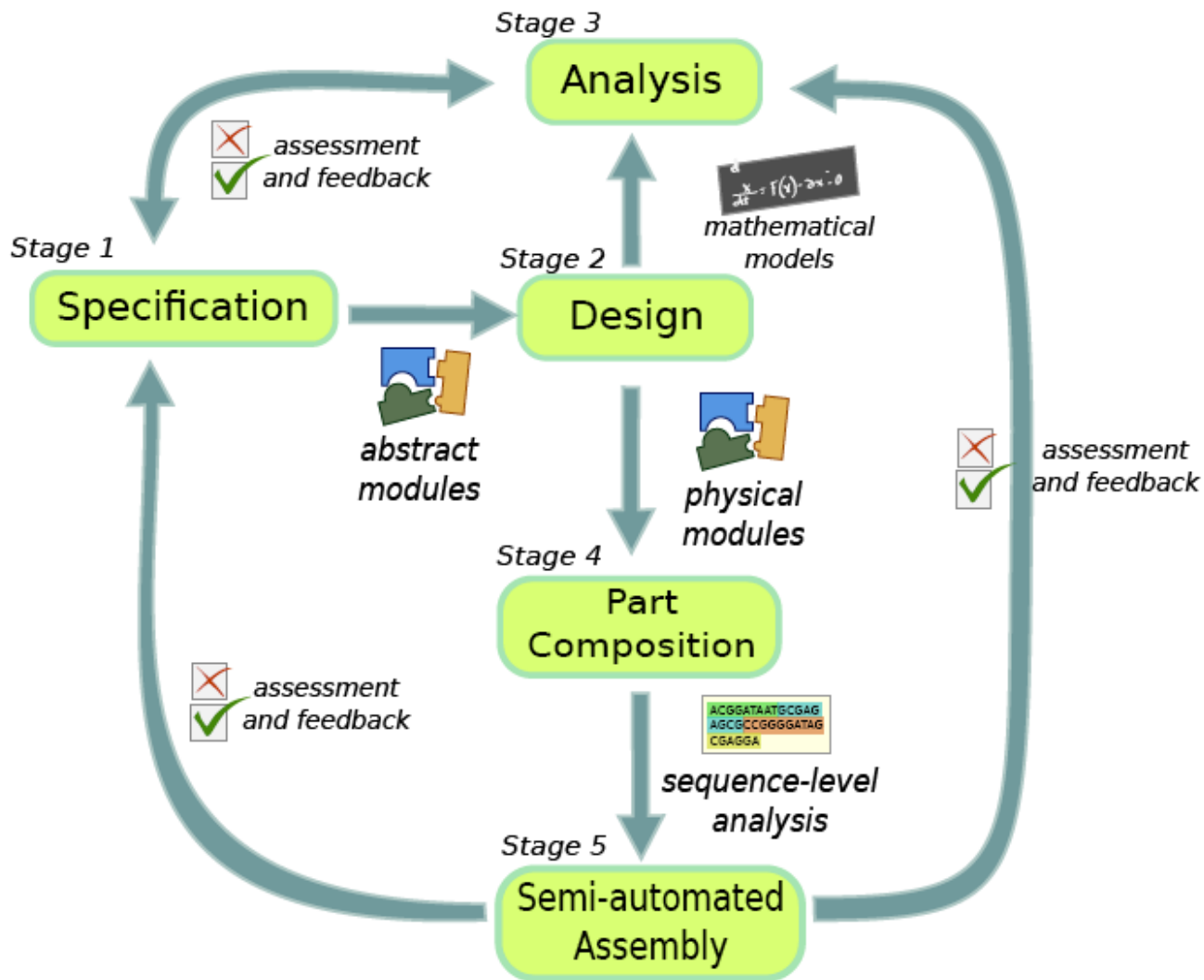
Timothy S. Gardner\*†, Charles R. Cantor\* & James J. Collins\*†

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

---

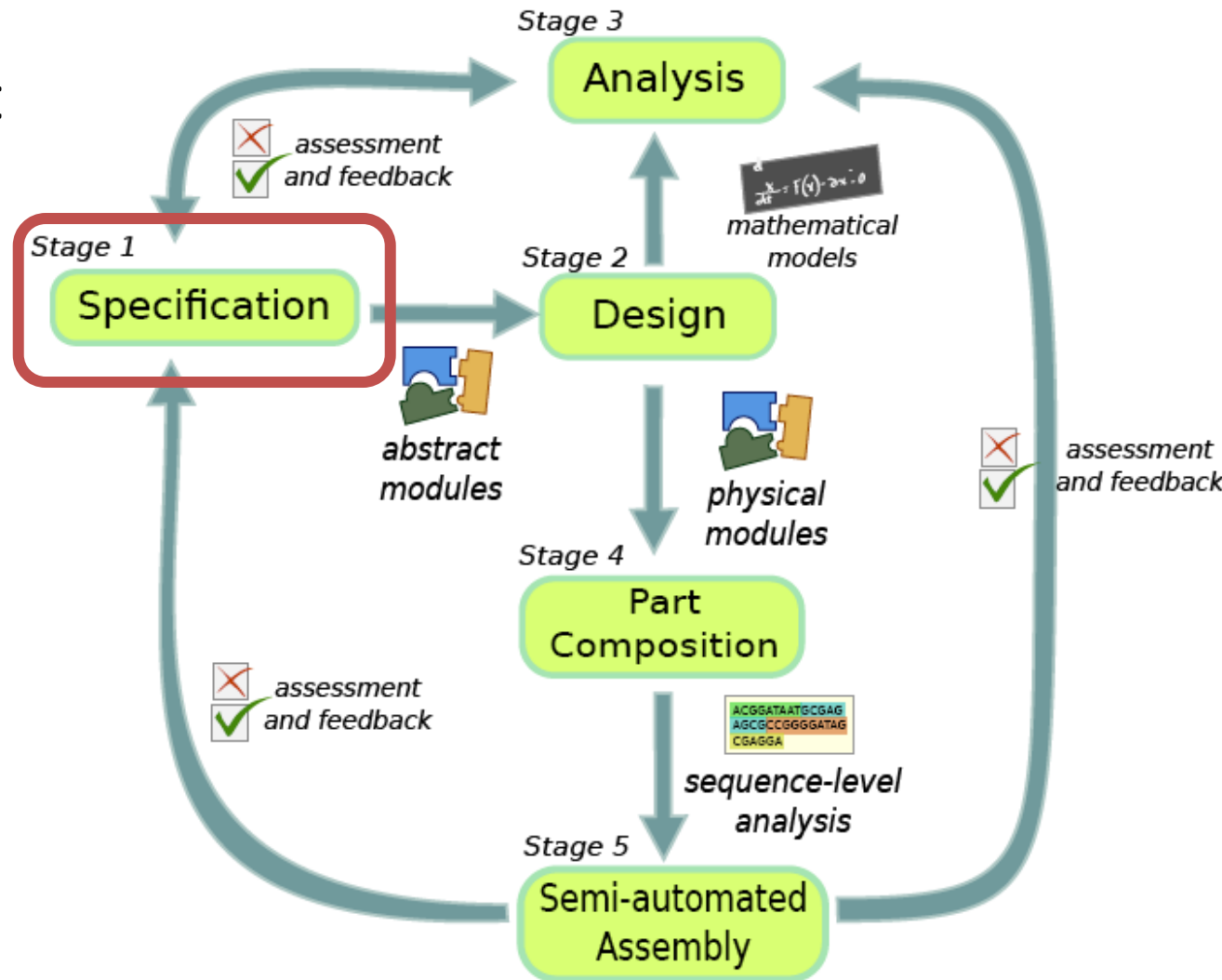
It has been proposed<sup>1</sup> that gene-regulatory circuits with virtually any desired property can be constructed from networks of simple regulatory elements. These properties, which include multistability and oscillations, have been found in specialized gene circuits such as the bacteriophage  $\lambda$  switch<sup>2</sup> and the Cyanobacteria circadian oscillator<sup>3</sup>. 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



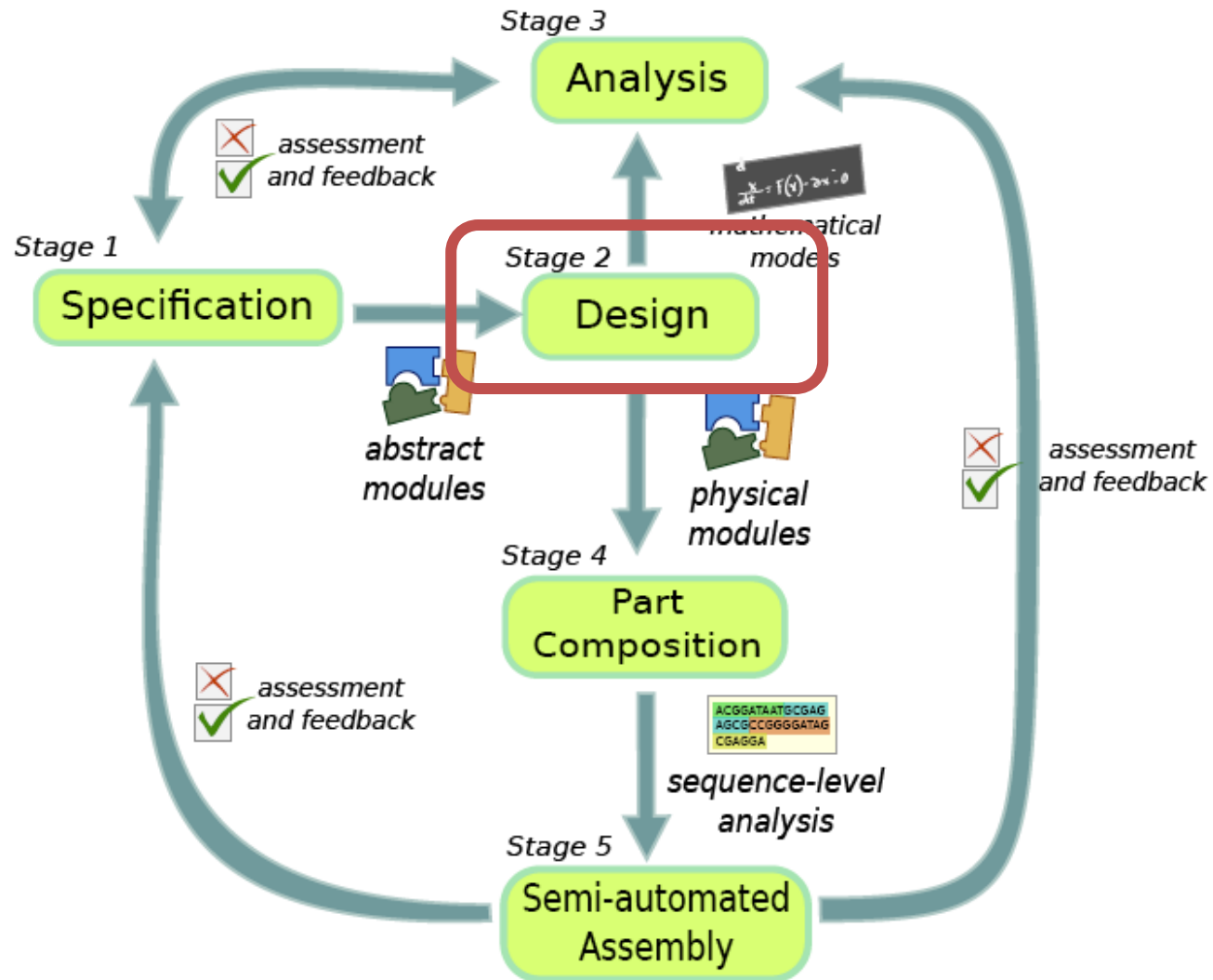
# Specification

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



# Design

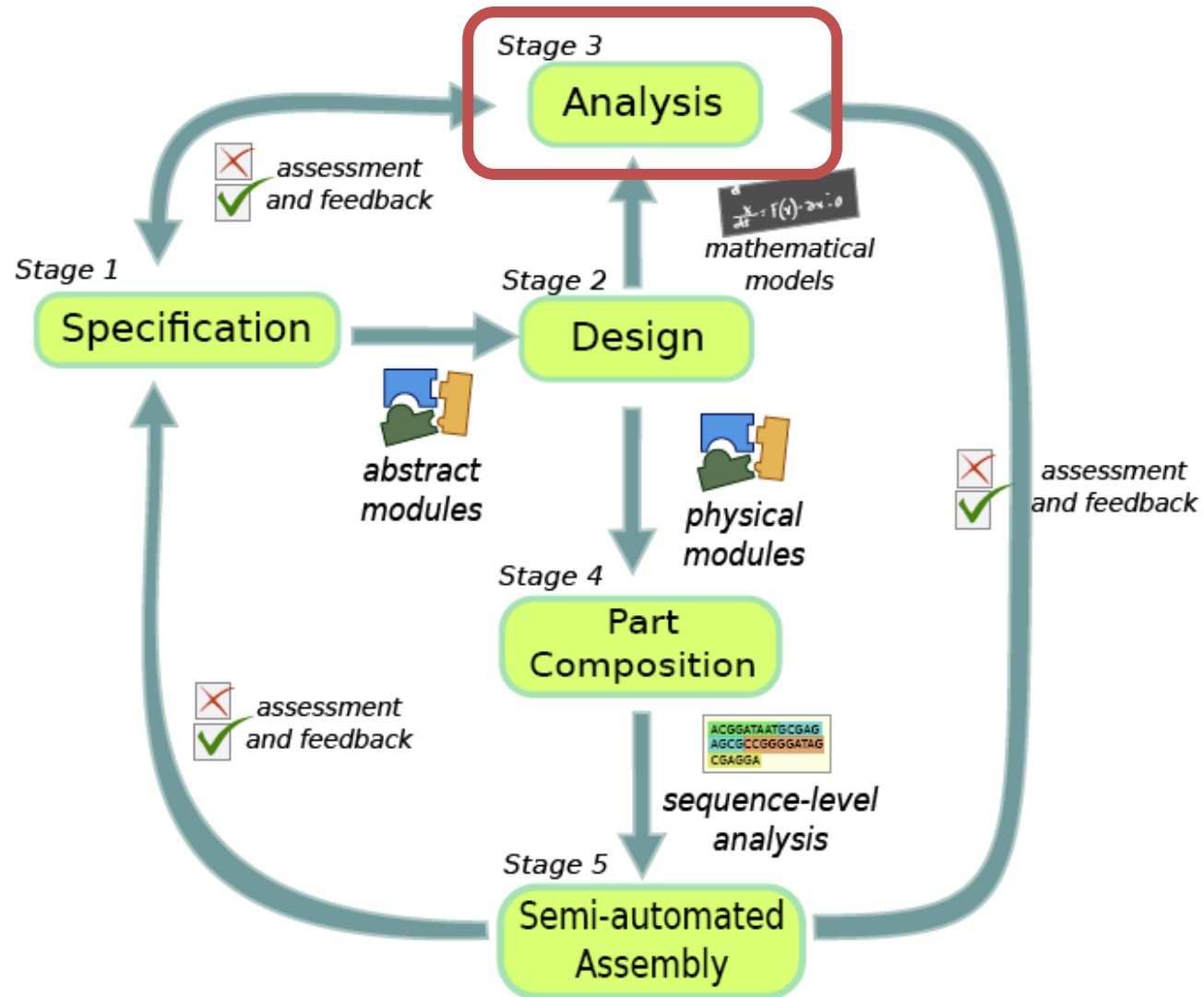
Provide the engineer with a set of building blocks that ease the task of making complex systems while ensuring that the specification is adhered to.



# Analysis

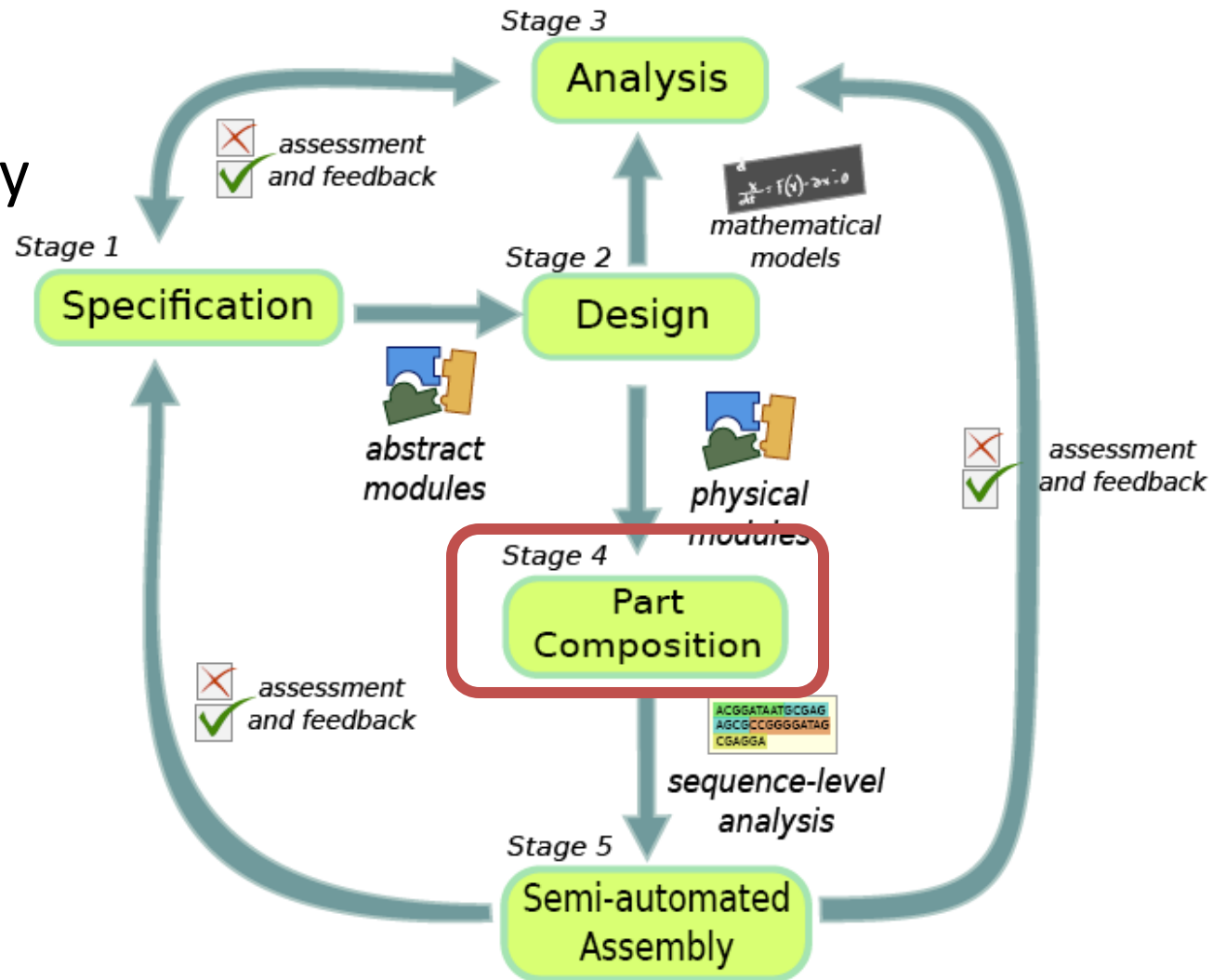
SBML/CellML

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

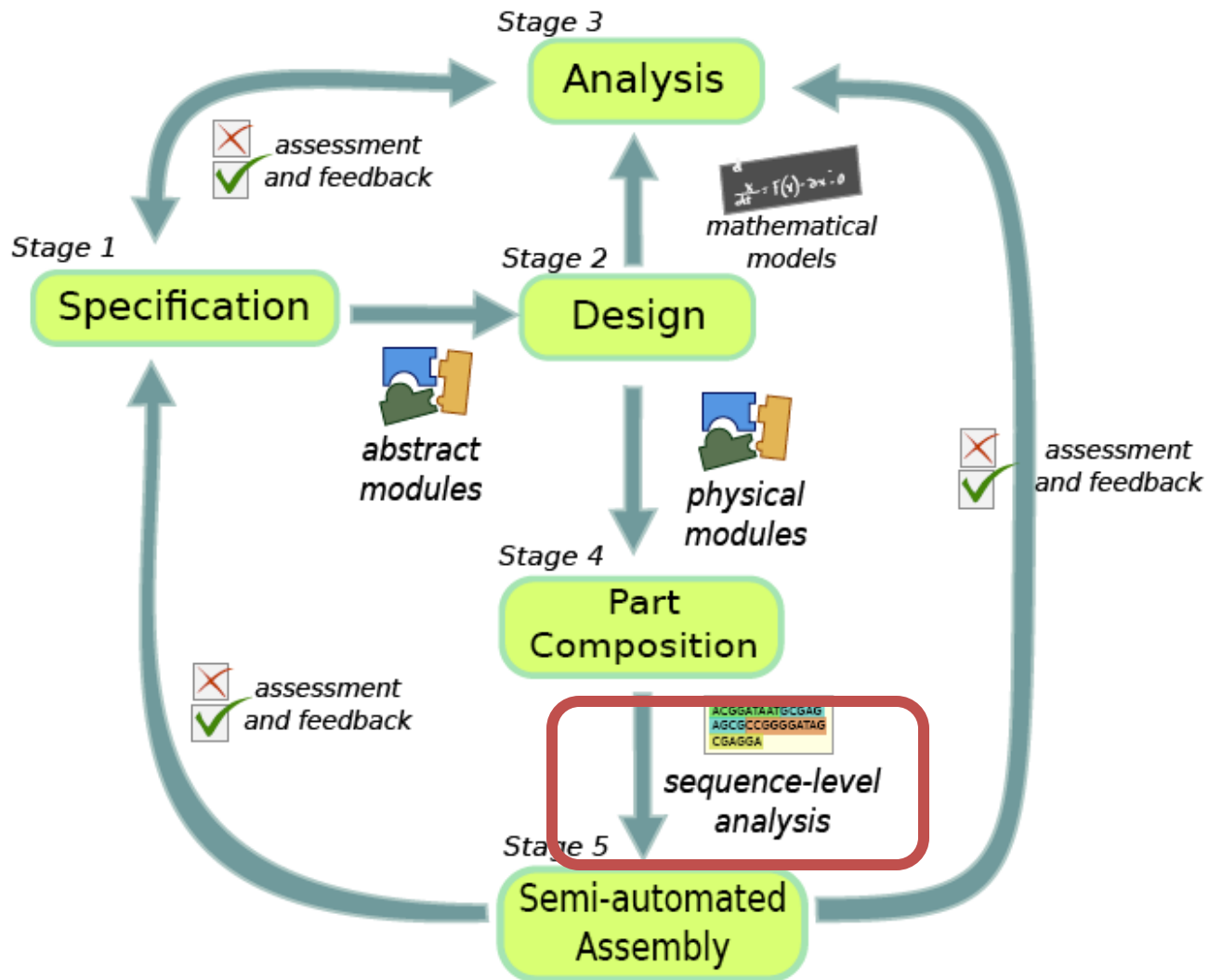


# Part Composition

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



# Computer Aided Design in Synthetic Biology

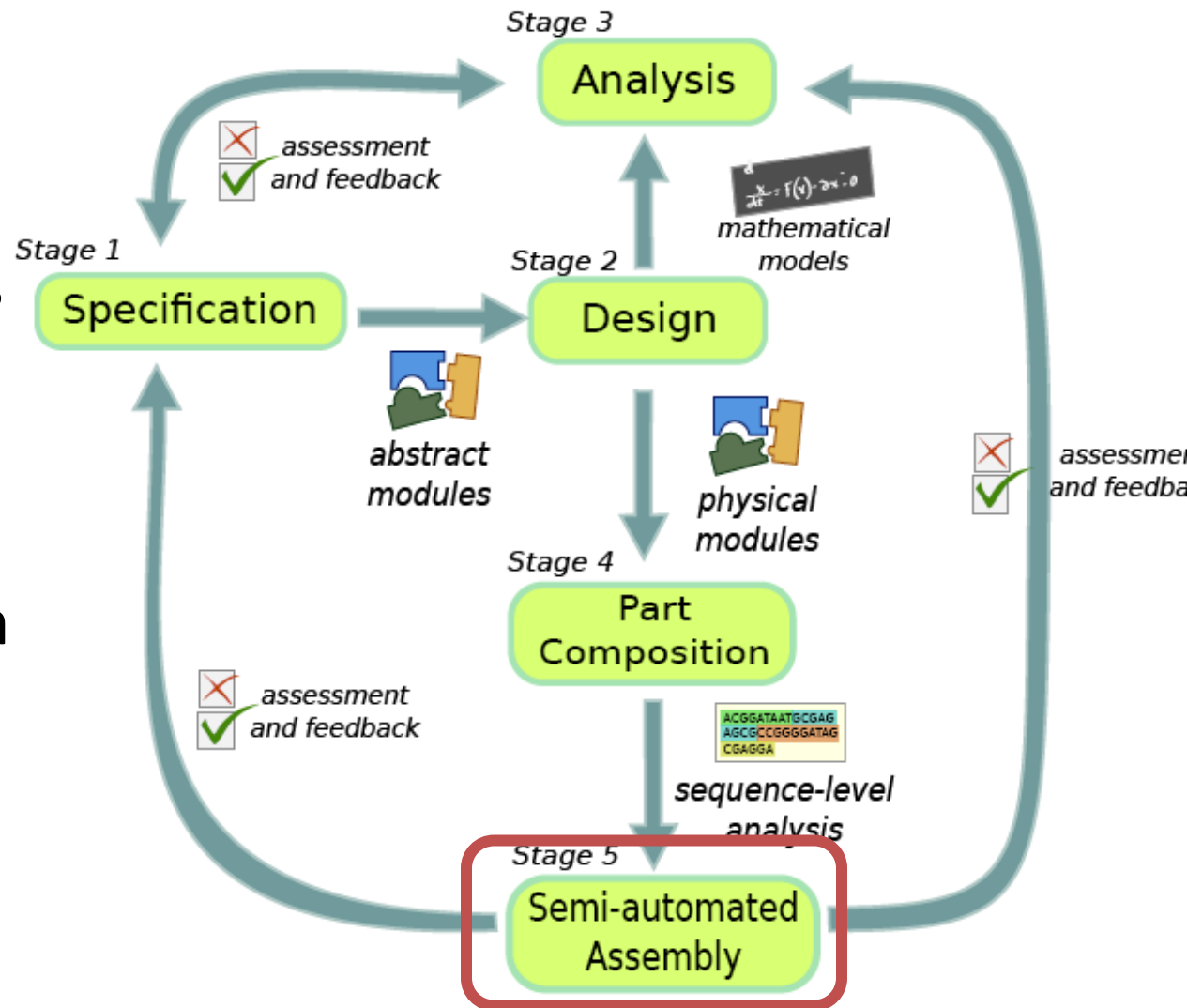


# Sequence Refinement

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

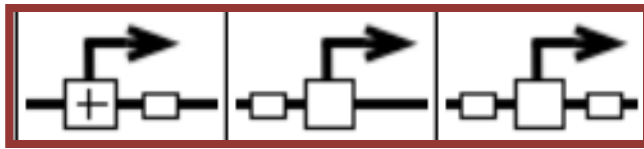
# Overall Aim of the Standardization Effort

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

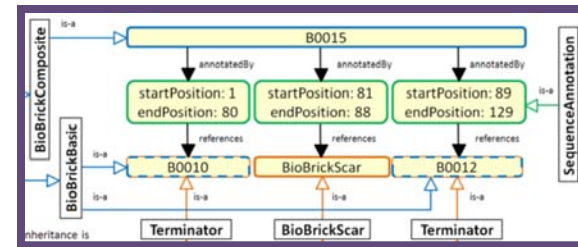
## Synthetic Biology Open Language (SBOL)



SBOL-visual

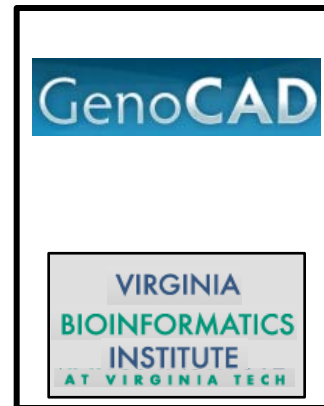


biofab



SBOL-semantic

UNIVERSITY of  
WASHINGTON

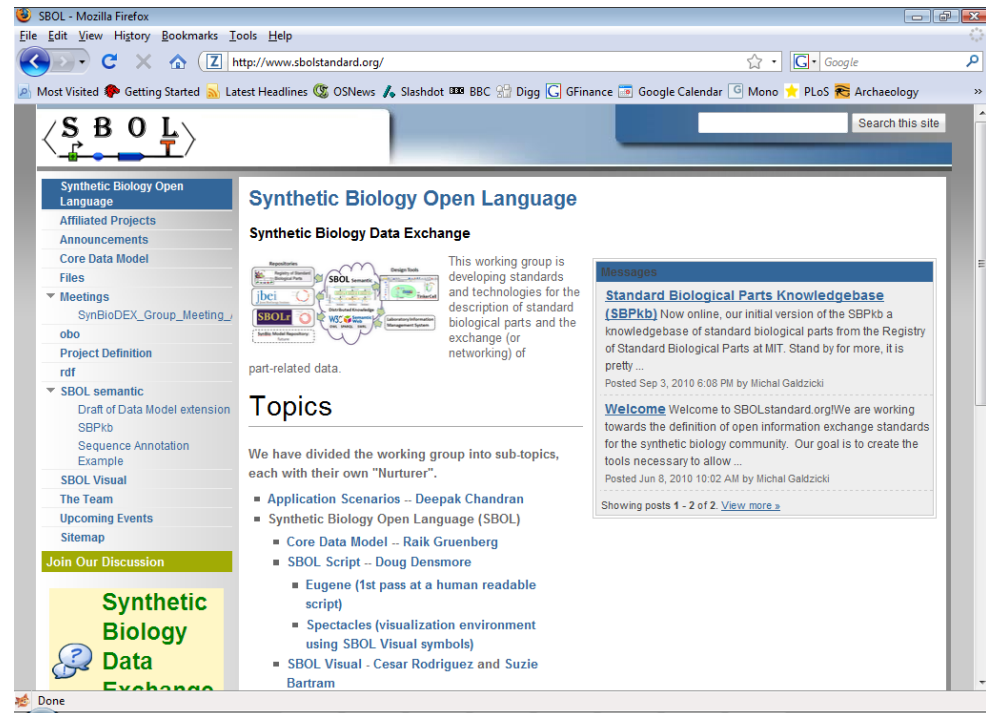


# Synthetic Biology Open Language: SBOL

<http://www.sbolstandard.org/>

1. SBOL Semantic
2. SBOL Visual

Eugene Scripting  
from Douglas  
Densmore (BU)



The screenshot shows the SBOL website in a Mozilla Firefox browser window. The address bar displays <http://www.sbolstandard.org/>. The page features a navigation menu on the left with categories such as "Synthetic Biology Open Language", "Affiliated Projects", "Announcements", "Core Data Model", "Files", "Meetings", "obo", "Project Definition", "rdf", "SBOL semantic", "SBOL Visual", "The Team", "Upcoming Events", and "Sitemap". The main content area is titled "Synthetic Biology Open Language" and includes a section for "Synthetic Biology Data Exchange" with a diagram of the SBOL ecosystem. Below this is a "Topics" section listing sub-topics like "Application Scenarios", "Synthetic Biology Open Language (SBOL)", "Core Data Model", "SBOL Script", "Eugene", "Spectacles", and "SBOL Visual". A "Messages" section on the right contains two posts, one dated Sep 3, 2010, and another dated Jun 8, 2010. A "Synthetic Biology Data Exchange" banner is visible at the bottom left of the page.

# SBOL Visual (SBOLv)



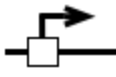

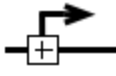

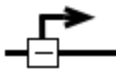
Developed at the Berkeley BIOFAB (Endy and Arkin)

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

# SBOL Visual (Implemented in Spectacles and TinkerCell)

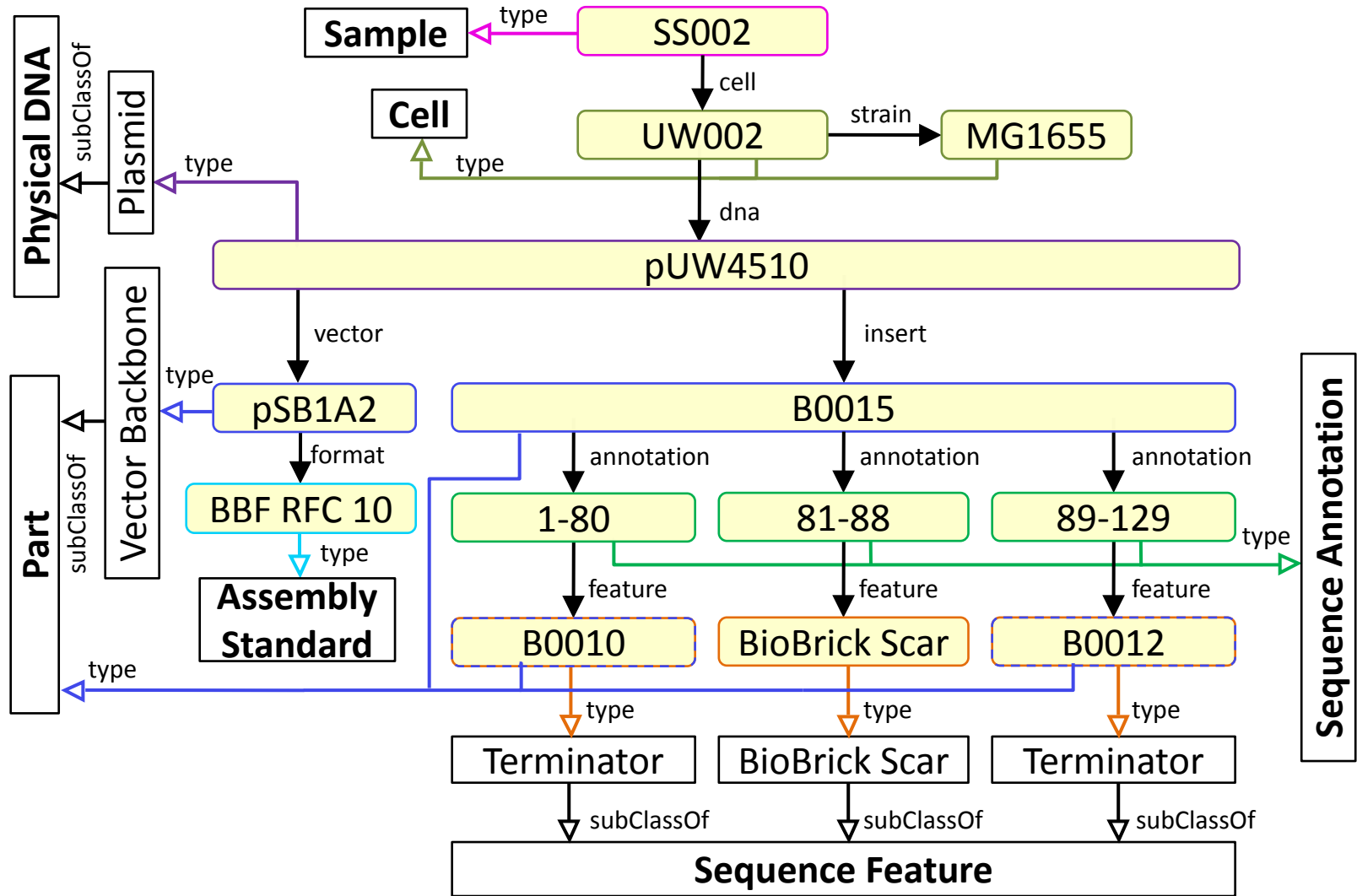
The screenshot displays the TinkerCell software interface for synthetic biology CAD. The main workspace shows a genetic circuit diagram with two identical units connected by a feedback loop. Each unit consists of a plasmid with an origin of replication (o1, o2), an inducer (i1), a ribosome binding site (rbs1), and a coding region (c1). The circuit includes various regulatory elements like promoters (P1, P2), operators (o1, o2, o3, o4), and ribosomes (rbs1, rbs2). Junctions (J1-J11) and RNA molecules (rna1, rna2) are also shown.

The right-hand side of the interface features a 'Tools Window' with a 'History' section and a 'Model summary' table. The table lists the components and their attributes and values:

Name	Attribute	Value
P1	concentration	1
P2	concentration	1
rna1	concentration	1
rna2	concentration	1
c1	activity	1
c2	activity	1
i1	activity	1
i2	activity	1
o1	activity	1
o2	activity	1
o3	activity	1
o4	activity	1
rbs1	activity	1

Below the table is a 'Console Window' displaying the following text:

```
*****  
***** (supplied with this  
release) for details  
*****  
*****  
*****  
Python modules loaded: Numpy,  
Scipy, PySCeS, NetworkX. Edit init.py  
to add more.  
>>>
```



**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("ATCCCC"));
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)**

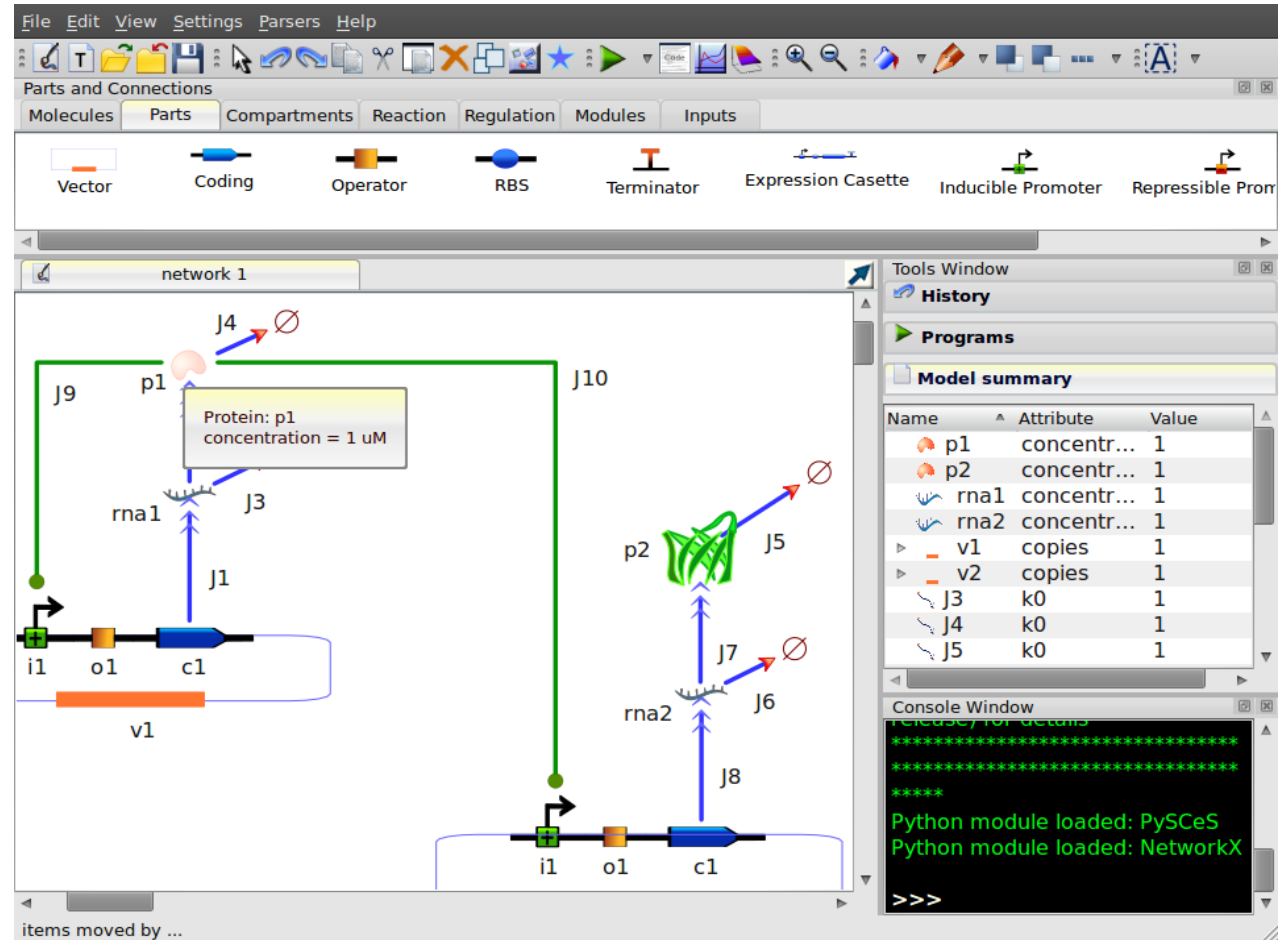
Test parts repository, ~5000 parts

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

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

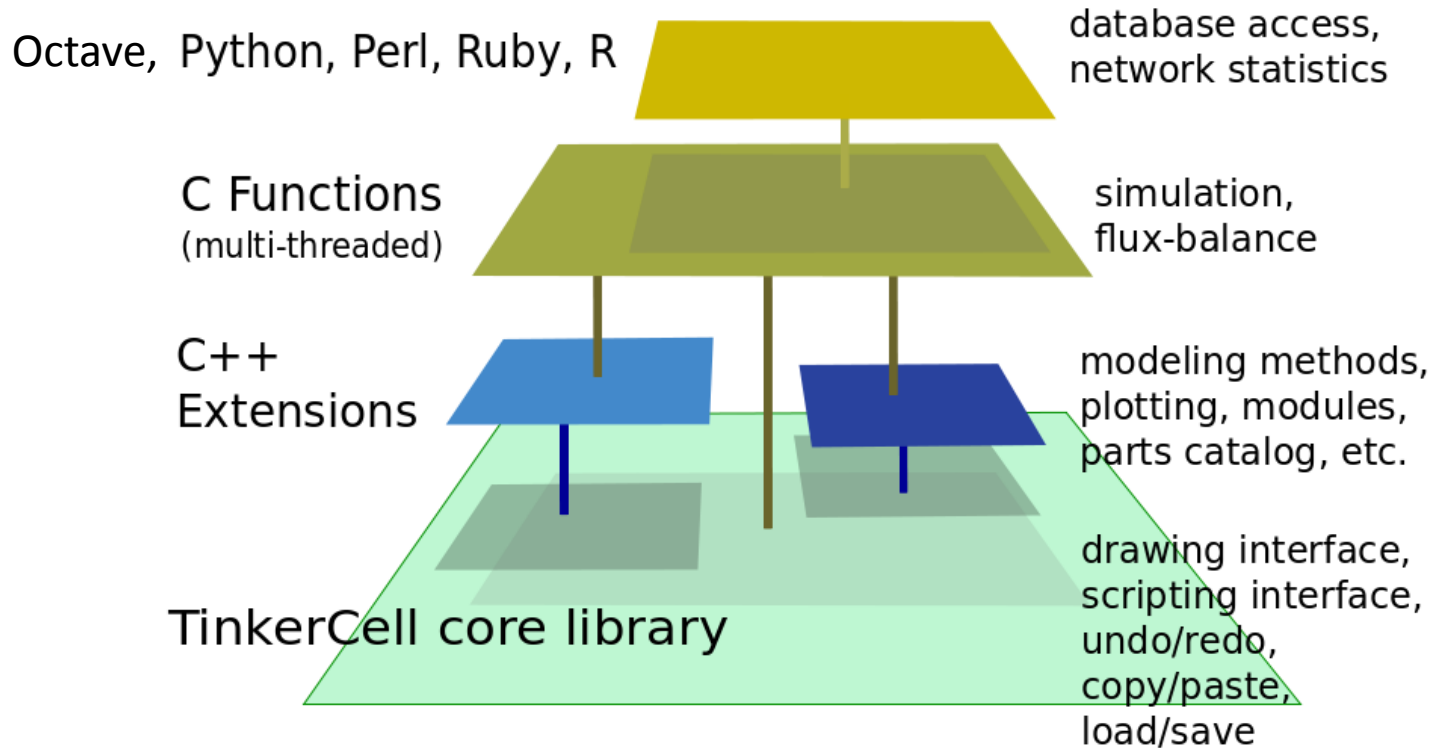
# TinkerCell: Project to explore the potential of computer aided design in synthetic biology

First prototype called Athena developed by Bergmann and Chandran

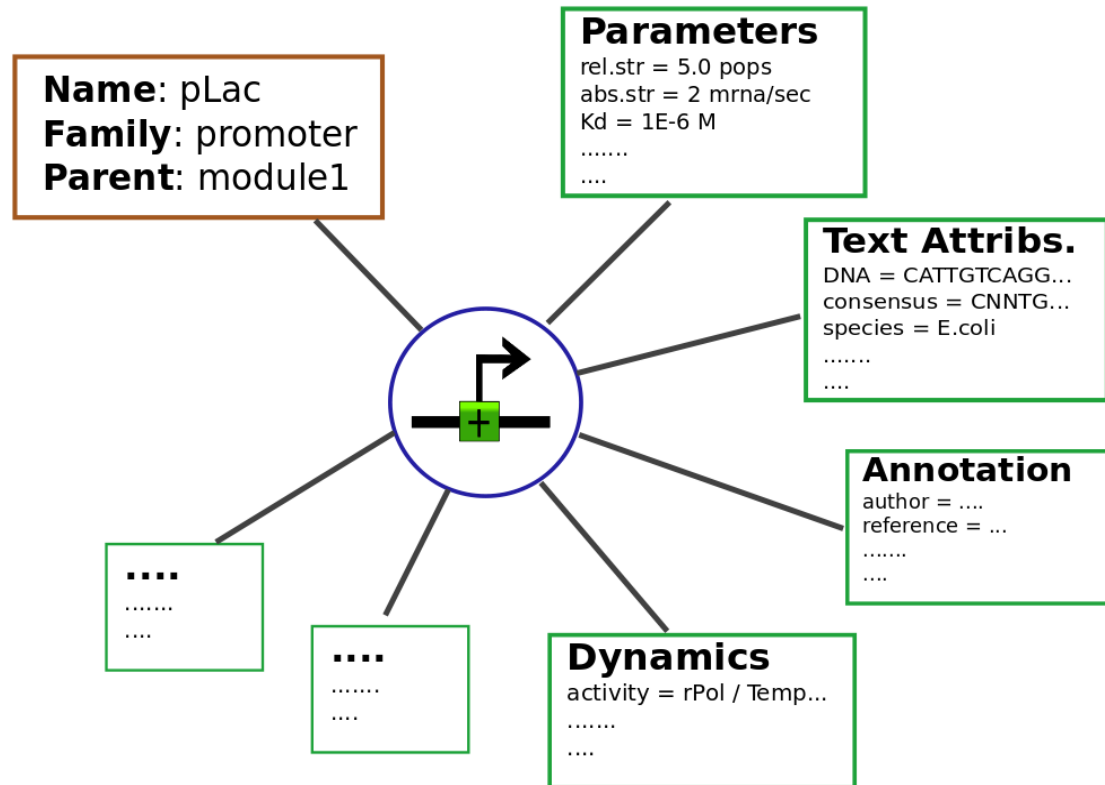


# Layered Architecture: Based on C++/Qt

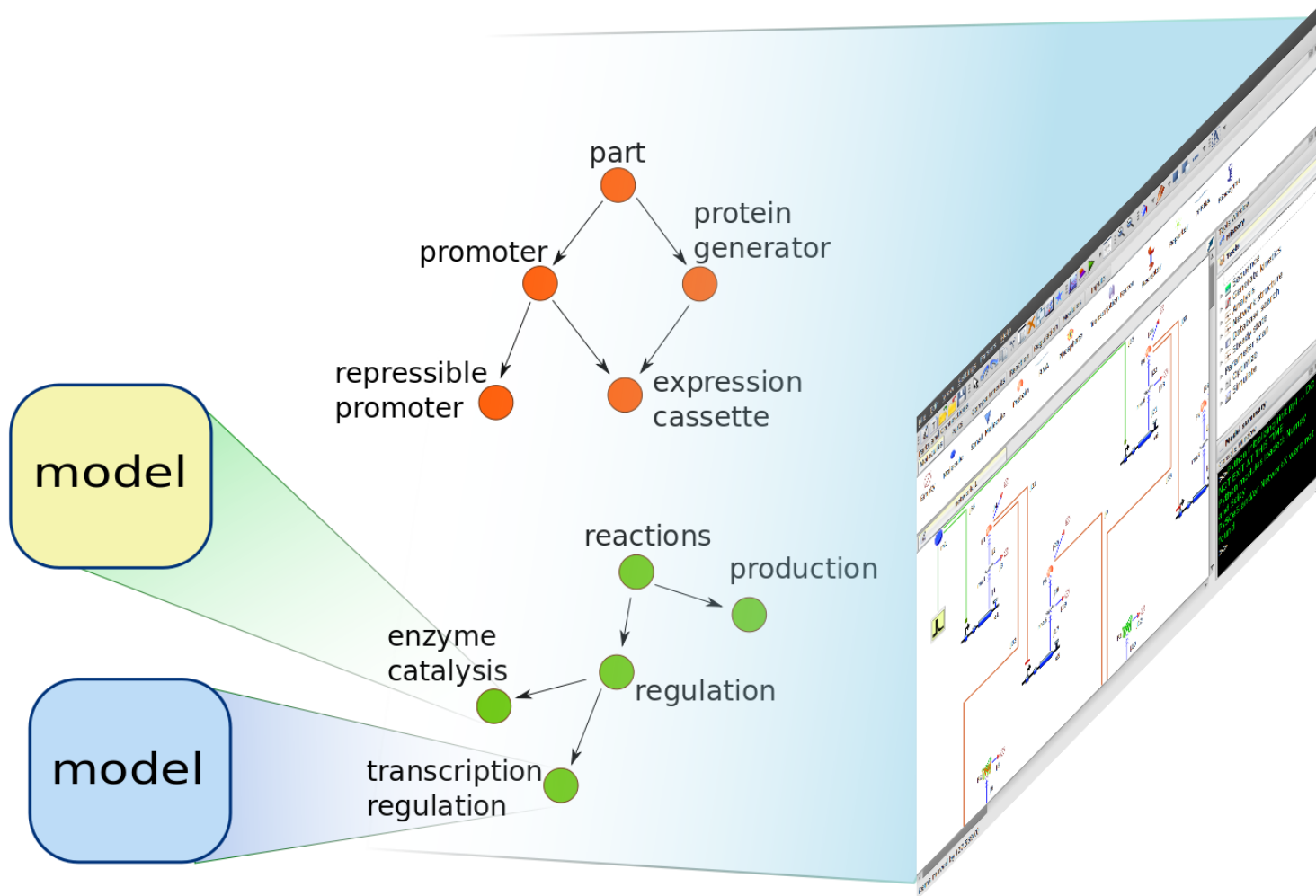
example



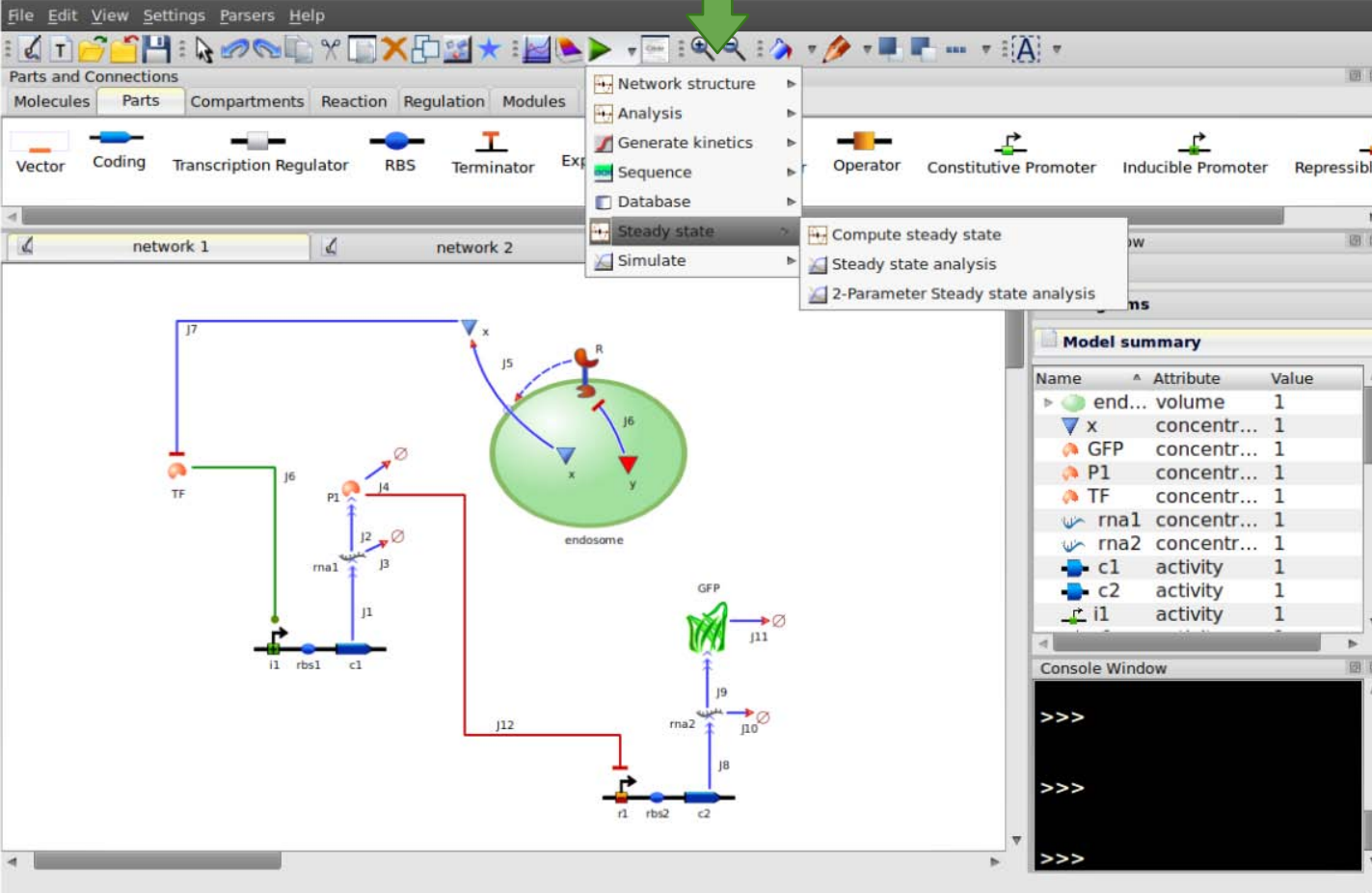
# Each component in the TinkerCell diagram is associated with one or more tables



# TinkerCell supports hierarchical sub-models



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



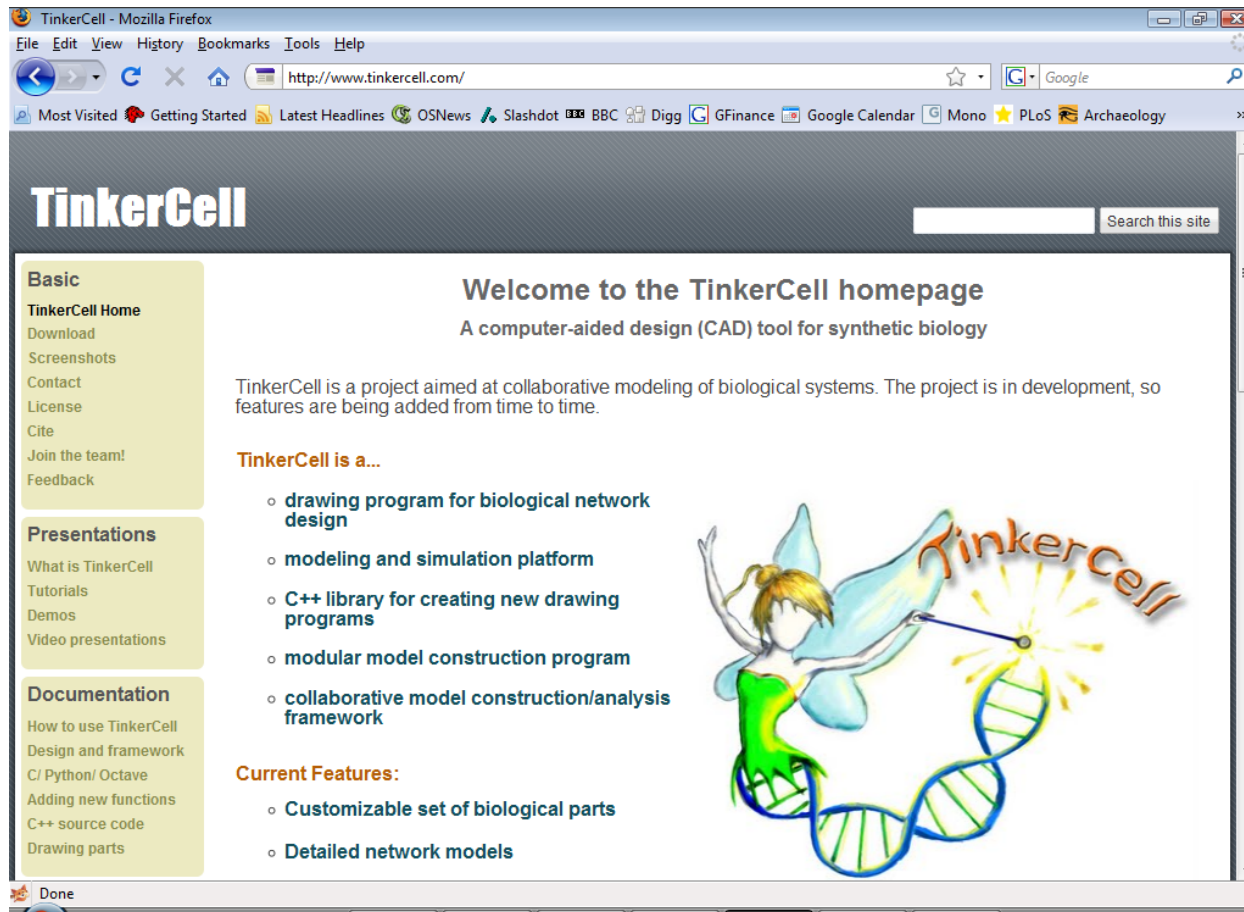
The screenshot displays the CellDesigner software interface. The main window shows a biological network diagram with various components like transcription factors (TF), ribosomes (rbs), and endosomes. A green arrow points to the 'Simulate' option in the 'Analysis' menu. The 'Steady state' sub-menu is open, showing options for 'Compute steady state', 'Steady state analysis', and '2-Parameter Steady state analysis'. The 'Model summary' table is visible on the right, listing model components and their values.

Name	Attribute	Value
end...	volume	1
x	concentr...	1
GFP	concentr...	1
P1	concentr...	1
TF	concentr...	1
rna1	concentr...	1
rna2	concentr...	1
c1	activity	1
c2	activity	1
i1	activity	1

# Availability

[www.tinkercell.com](http://www.tinkercell.com) (Windows, Mac and Linux, released under BSD)

Contact author for details (dchandran1@gmail.com)



The screenshot shows a Mozilla Firefox browser window displaying the TinkerCell homepage. The browser's address bar shows the URL <http://www.tinkercell.com/>. The page features a dark header with the TinkerCell logo and a search bar. The main content area is divided into a left sidebar and a central main section. The sidebar contains three categories: 'Basic' (with links like TinkerCell Home, Download, Screenshots, Contact, License, Cite, Join the team!, Feedback), 'Presentations' (with links like What is TinkerCell, Tutorials, Demos, Video presentations), and 'Documentation' (with links like How to use TinkerCell, Design and framework, C/Python/ Octave, Adding new functions, C++ source code, Drawing parts). The main section has a heading 'Welcome to the TinkerCell homepage' and a sub-heading 'A computer-aided design (CAD) tool for synthetic biology'. Below this, a paragraph states: 'TinkerCell is a project aimed at collaborative modeling of biological systems. The project is in development, so features are being added from time to time.' A section titled 'TinkerCell is a...' lists several features: drawing program for biological network design, modeling and simulation platform, C++ library for creating new drawing programs, modular model construction program, collaborative model construction/analysis framework, and Current Features: Customizable set of biological parts, Detailed network models. To the right of the text is an illustration of a fairy with blue wings and a green dress, holding a glowing yellow wand that points to a glowing blue and green DNA double helix. The word 'TinkerCell' is written in a stylized, glowing orange font above the DNA.

**TinkerCell**

Search this site

## Welcome to the TinkerCell homepage

A computer-aided design (CAD) tool for synthetic biology


TinkerCell is a project aimed at collaborative modeling of biological systems. The project is in development, so features are being added from time to time.

**TinkerCell is a...**

- o drawing program for biological network design
- o modeling and simulation platform
- o C++ library for creating new drawing programs
- o modular model construction program
- o collaborative model construction/analysis framework

**Current Features:**

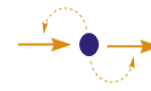
- o Customizable set of biological parts
- o Detailed network models



# C and Python packages included with TinkerCell

## Sundials

- Time-course simulation
- Steady state plot



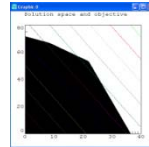
PySCeS

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



## SciPy

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

## Ip\_solve

- flux balance analysis

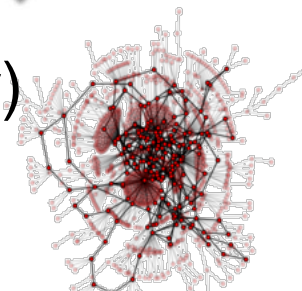


Antimony

SBML and CellML support

## NetworkX (lanl.gov)

- graph analysis
- graph layout



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

SBOLr

As of Saturday 12th June 2010  
there are 5169 Parts available

Welcome, Michal Galdzicki

search | Advanced Search | Blast

- Home
- News
- Entries
- My Profile
- Add new entry
- Update Account
- Feedback
- About Us
- Logout

		Most Recent	Most Voted	Most Commented							Stats	Created
#	Type	ID	Name	Summary	Owner							
1.	Part	UW_003503	BBa_S04137	heat sensitive cl QPI	Admin Semantic SBOL					0 0 0	01 Jan 2010	
2.	Part	UW_003504	BBa_J52028	BBa_J52028	Admin Semantic SBOL					1 0 1	01 Jan 2010	
3.	Part	UW_003505	BBa_S04055	Synthetic						0 0 0	01 Jan 2010	
4.	Part	UW_003501	BBa_P1014	gentamicin						0 0 0	01 Jan 2010	
5.	Part	UW_003502	BBa_I746220	AIP Detecto						0 0 0	01 Jan 2010	
6.	Part	UW_003499	BBa_K143056	Promoter xy						0 0 0	01 Jan 2010	
12.	Part	UW_003495	BBa_I14033	P(Cat)						0 0 0	01 Jan 2010	

Per page: 15 30 45 All

Export as: Printable ( )

BBa\_I14033

-35 -10

1 GGCACGTAAGAGGTTCCAACCTTTCCACCATAATGAAACA  
CCGTGCATTCTCCAAGGTTGAAAGTGGTATTACTTTGT

© University of Washington  
All rights reserved.  
Powered by: SBOL Semantic | Feedback | sbolr@uw.edu

SBOLr is the Sauro Lab branch of JBEI

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

# Demo

Building: 0.00  
Composition: 1.48



# Relationship to Efforts in the Systems Biology Community

Standards that may have utility in the synthetic  
biology field:

- SBML
- CellML
- SBGN
- SEDML
- BioPAX

# SBML

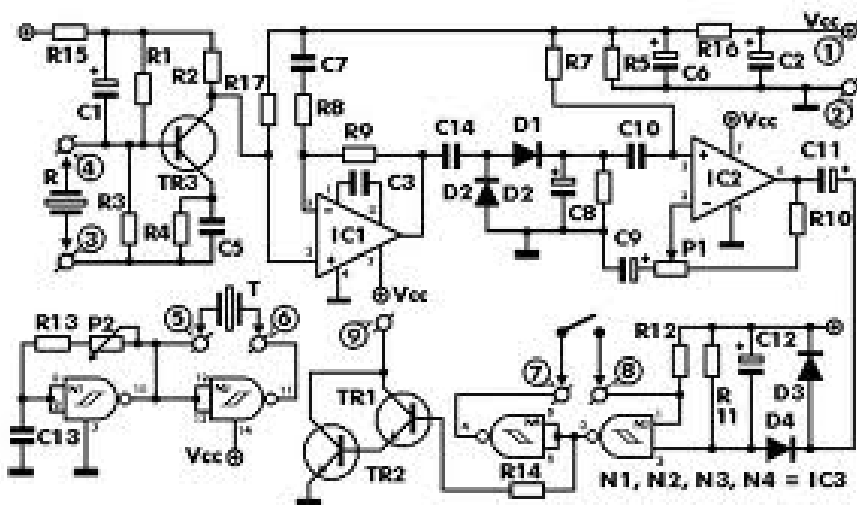
SBOL Semantic could be made into a SBML package, but no resource to currently undertake this task.

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

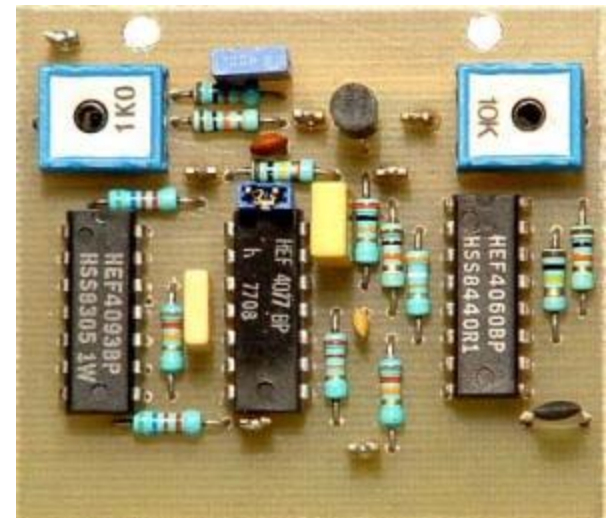
# SBGN

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.

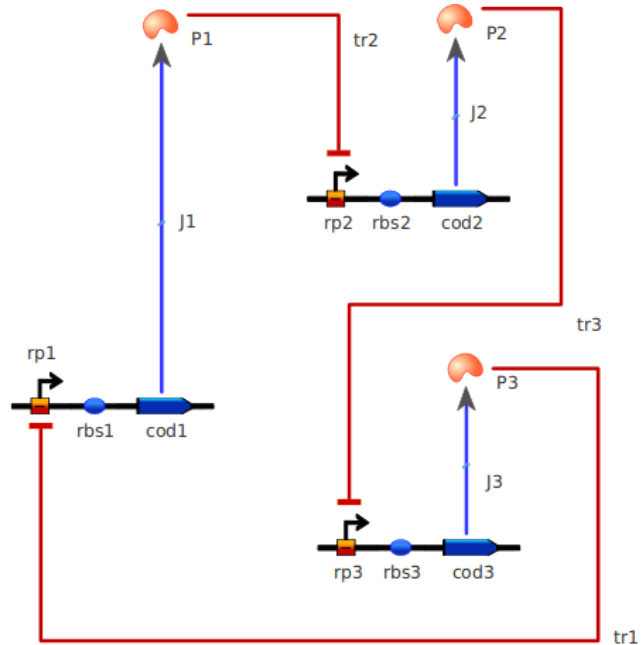
We need this:



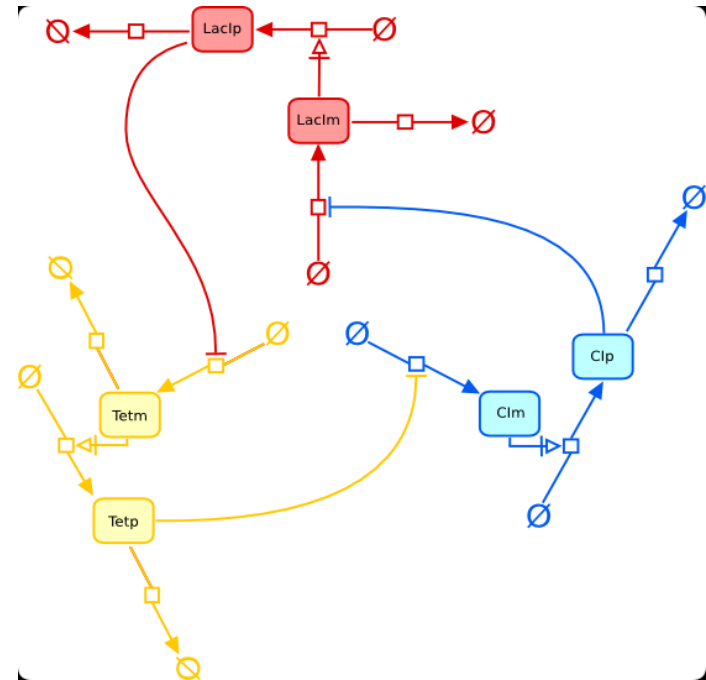
But SBGN is more like this:



# 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

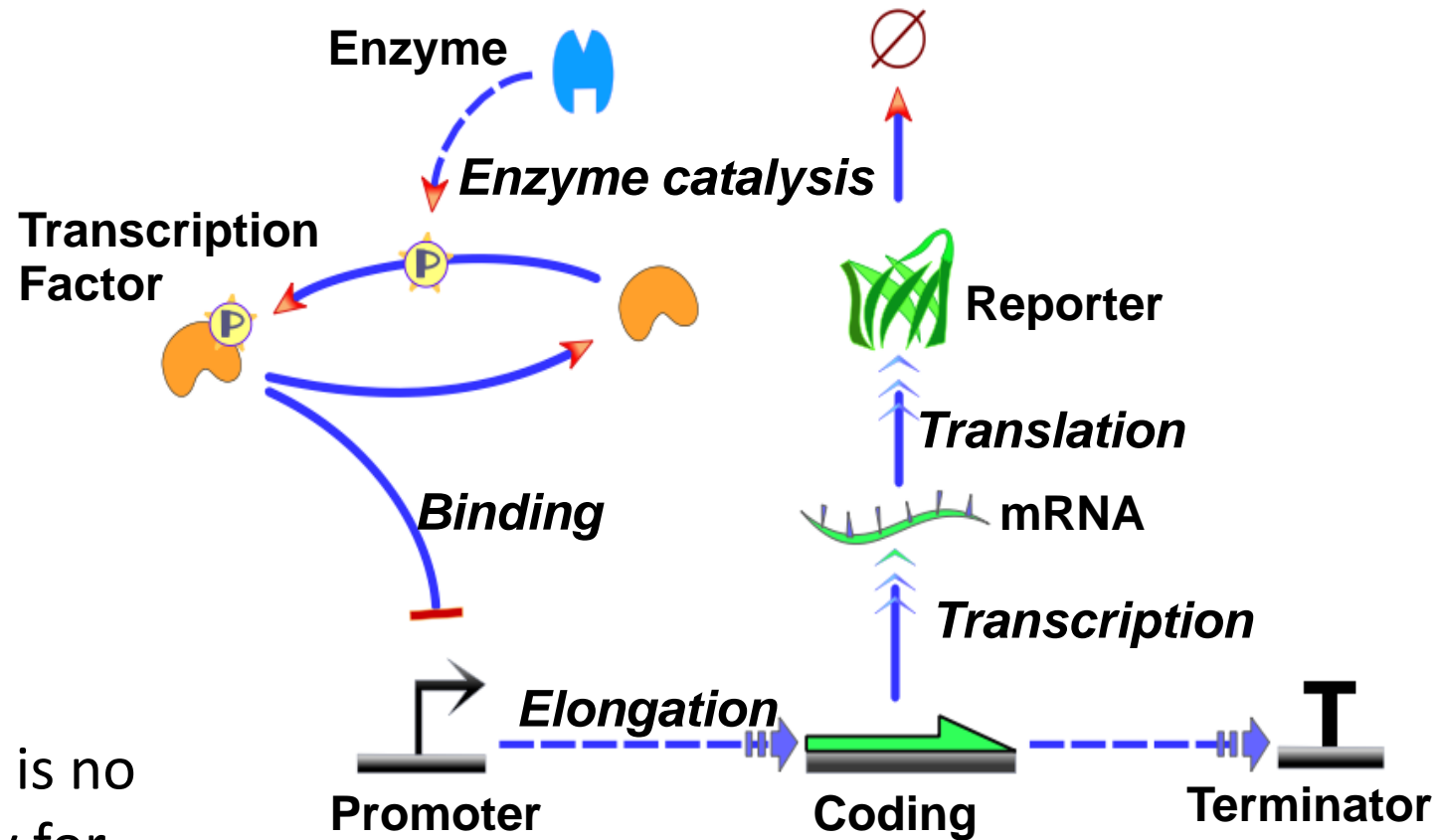
## **Funding**

National Library of Medicine (Galdziki)

Other funding scraped from NIH (JSIm/SBW) , NIH (SBW) and NSF (FIBR),

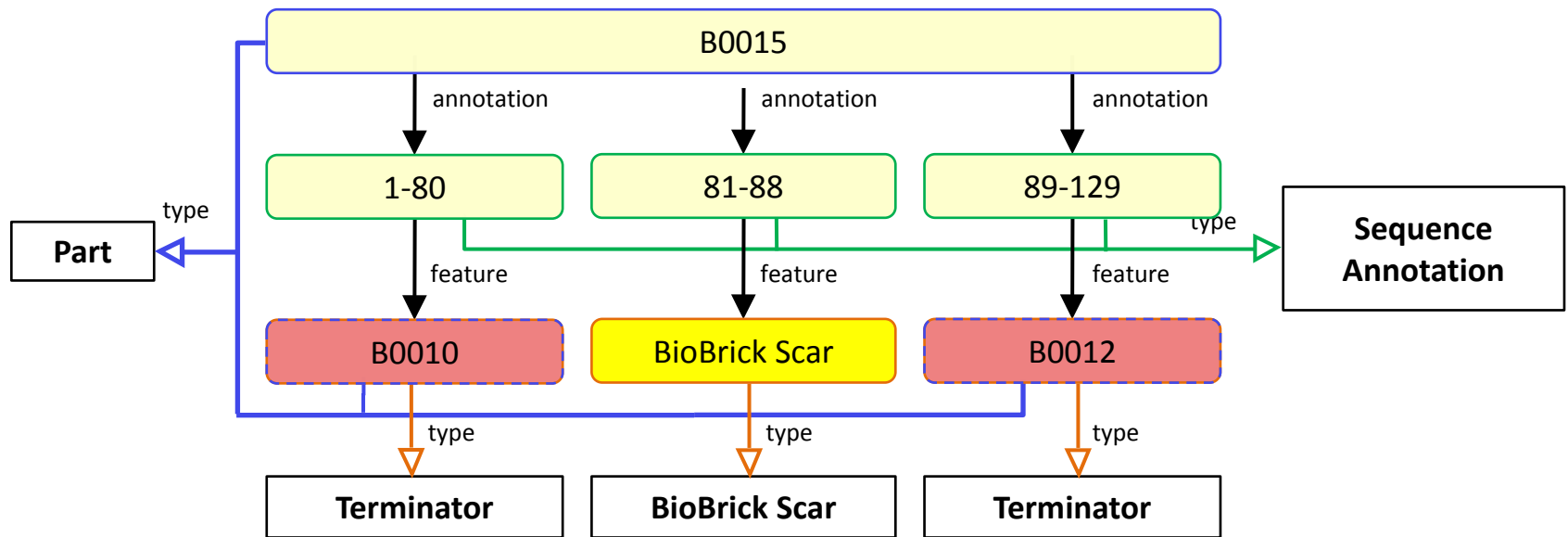
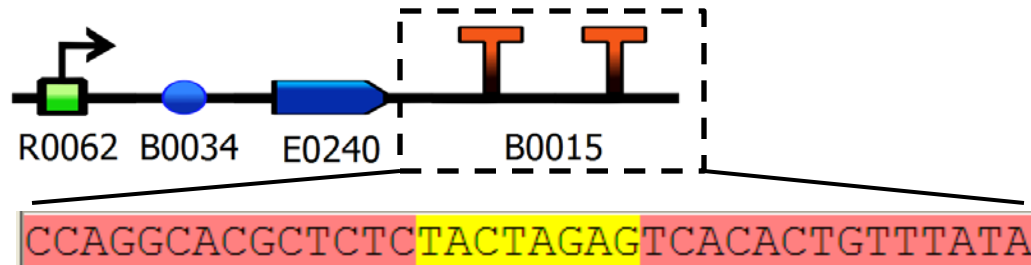
Microsoft Strategic Research, BioFAB, JBEI

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<sup>37</sup>