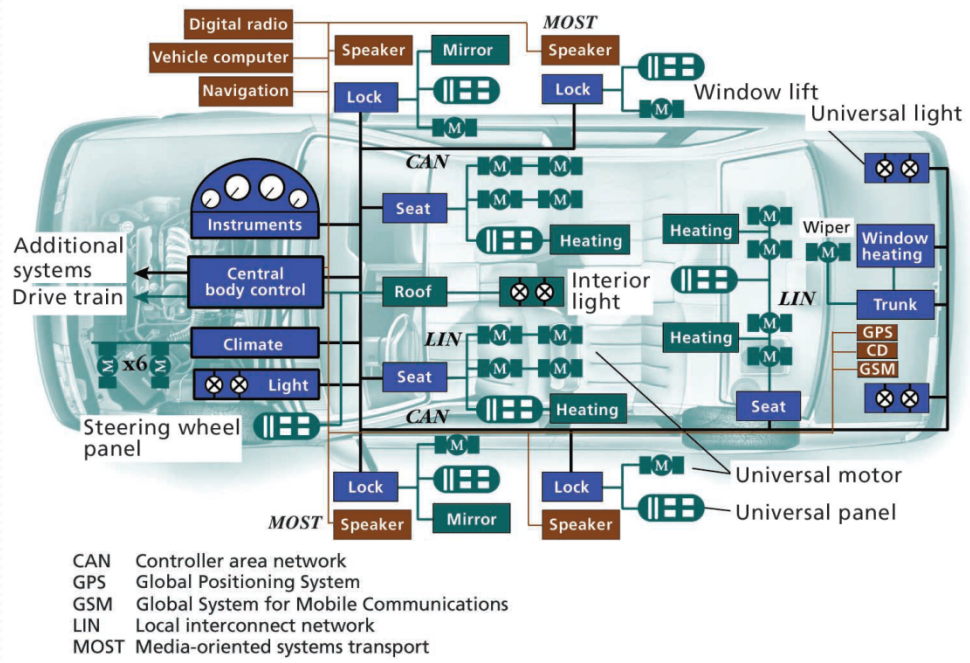
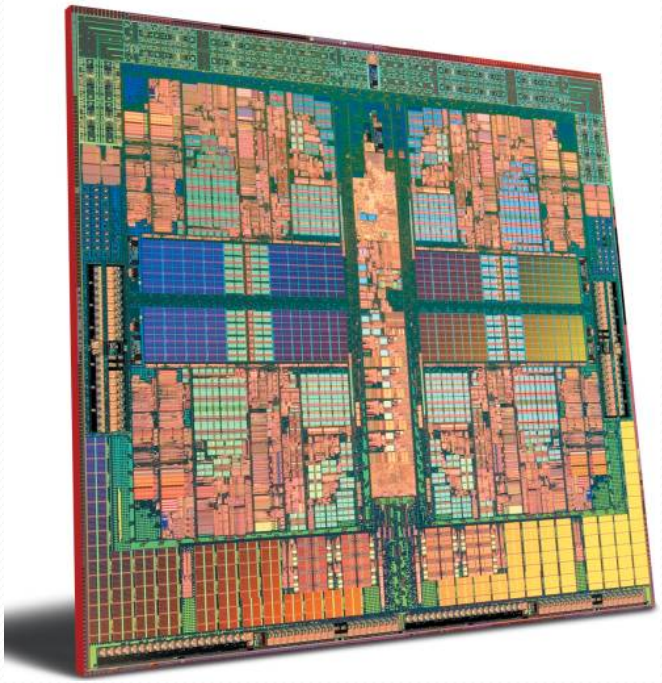
Design Automation in Synthetic Biology

Douglas Densmore

December 1, 2009



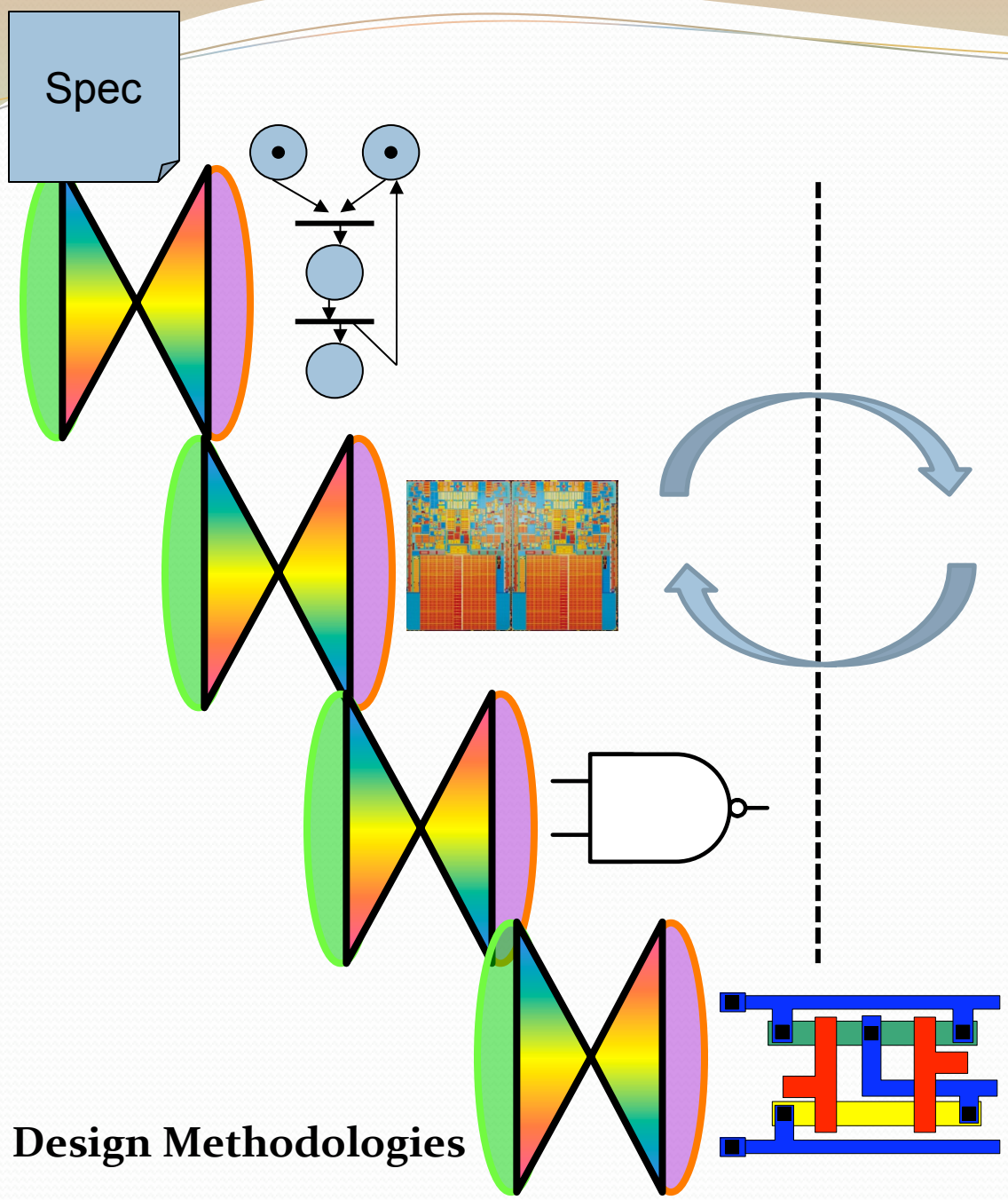
Heterogeneous Communication and Components



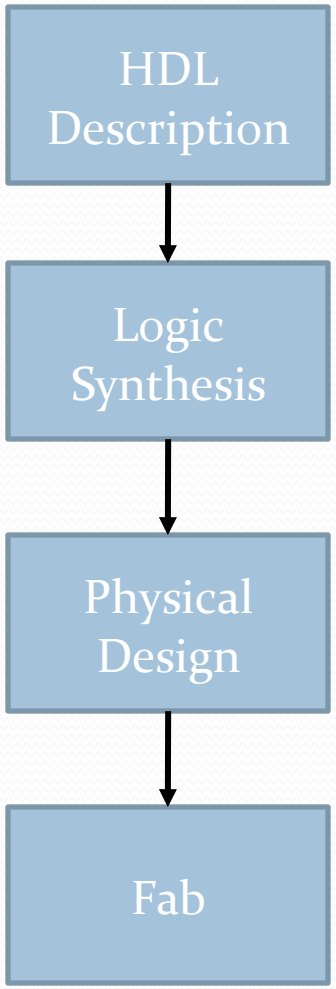
VLSI and Increased Parallelism

Four factors really drive electronic design automation: complexity, heterogeneity, time to market pressures, and deep sub-micron effects.

Need: Specification, design, and physical assembly tool flows

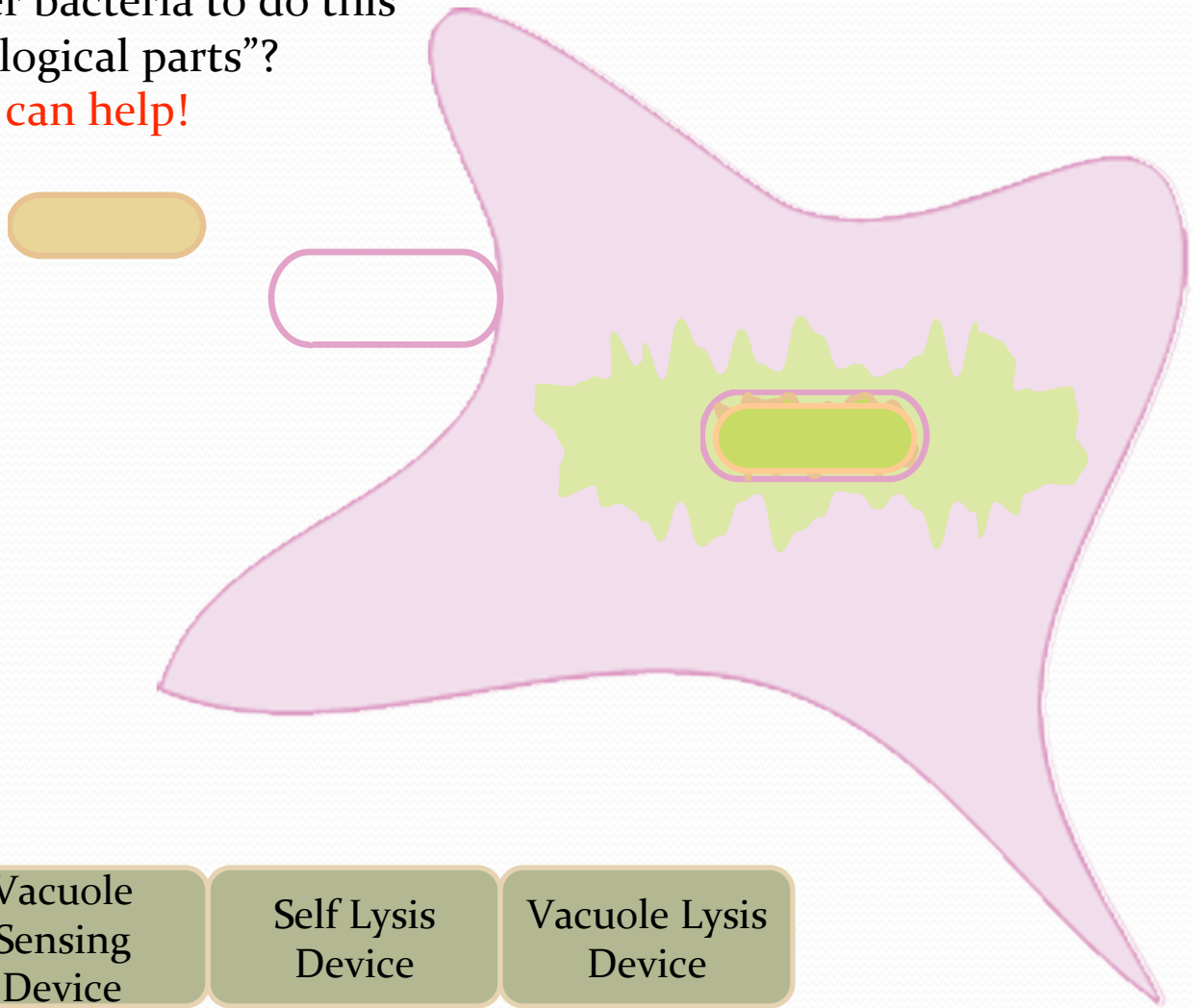


Design Methodologies



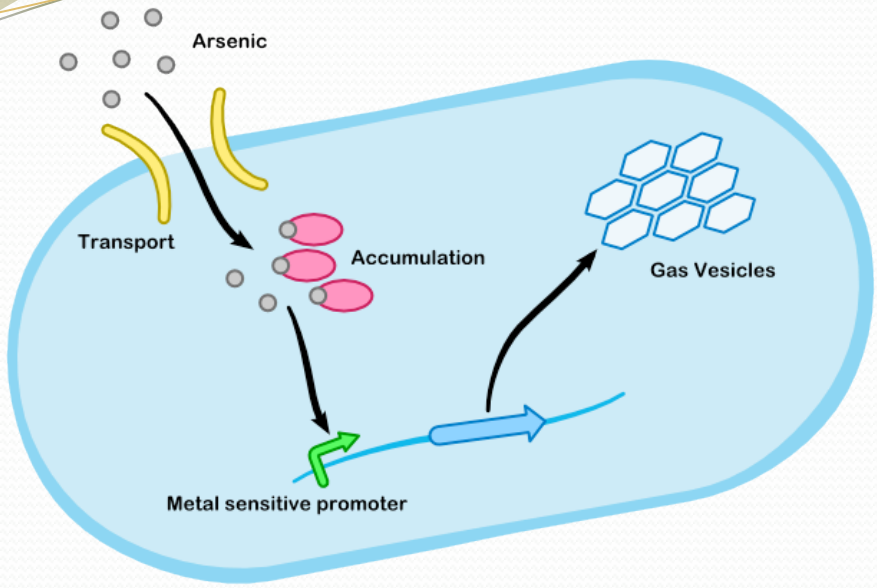
Design Flows

How do we engineer bacteria to do this using “standard biological parts”?
Design automation can help!

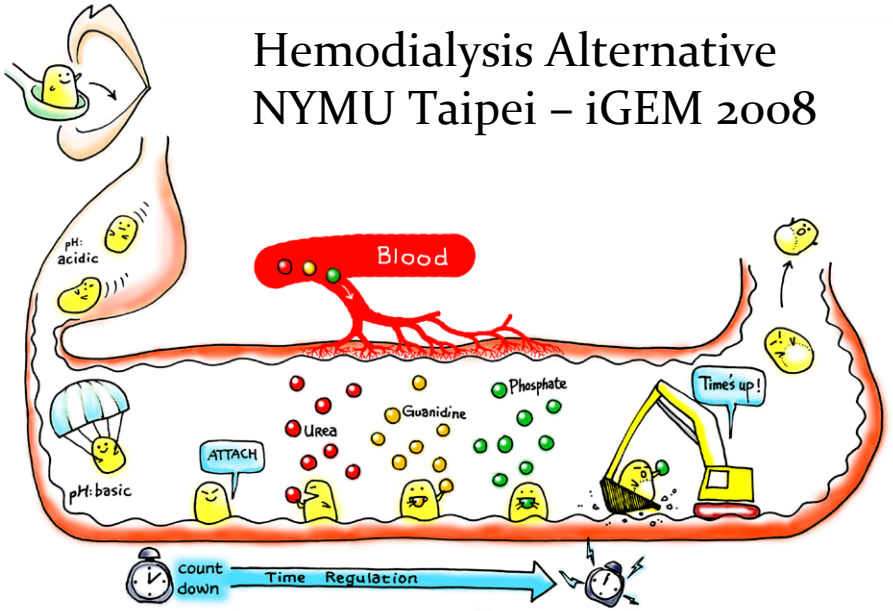


Engineered steps

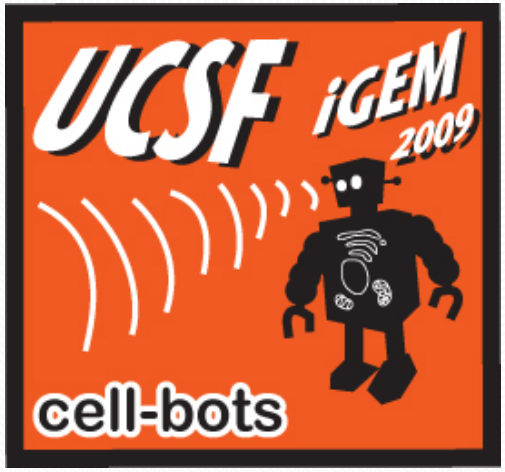
Invasion Device	Vacuole Sensing Device	Self Lysis Device	Vacuole Lysis Device



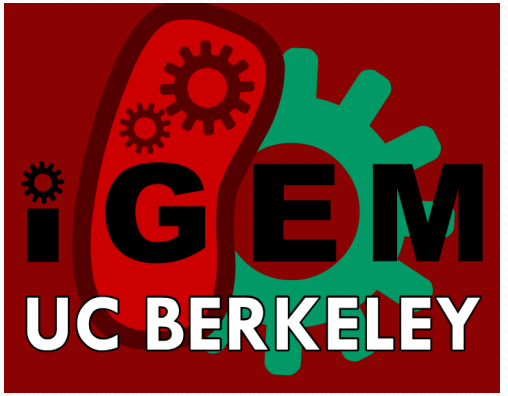
Arsenic Filtering System
Groningen – iGEM 2009



Hemodialysis Alternative
NYMU Taipei – iGEM 2008

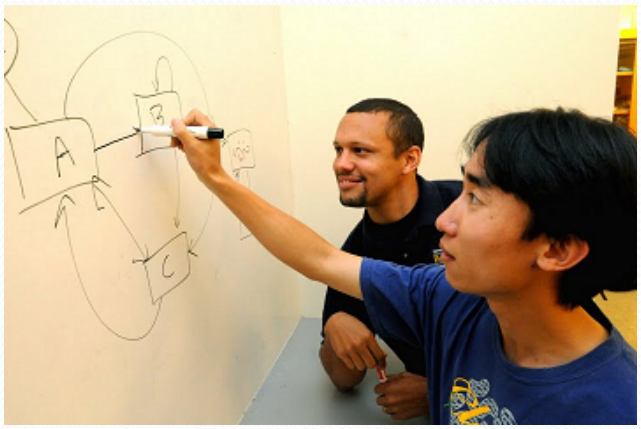


Chemotaxis control
and payload capture
UCSF – iGEM 2009

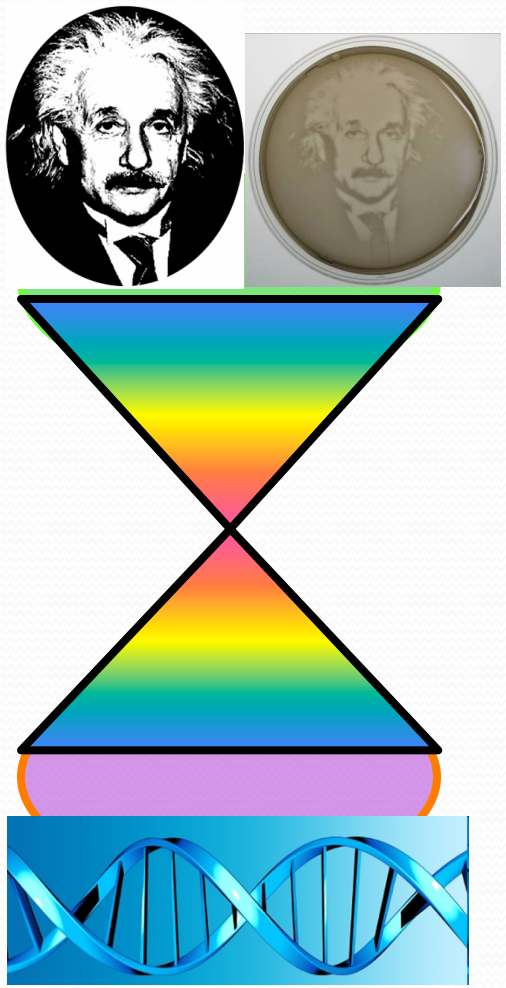


Red blood cell substitute
UC Berkeley – iGEM 2007

How are we going to design these biological systems in the future?

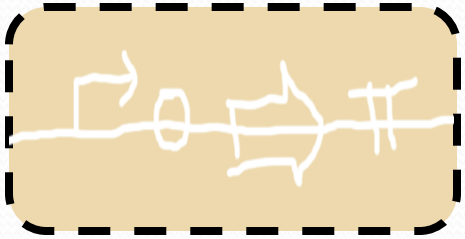

Design automation will play a key role

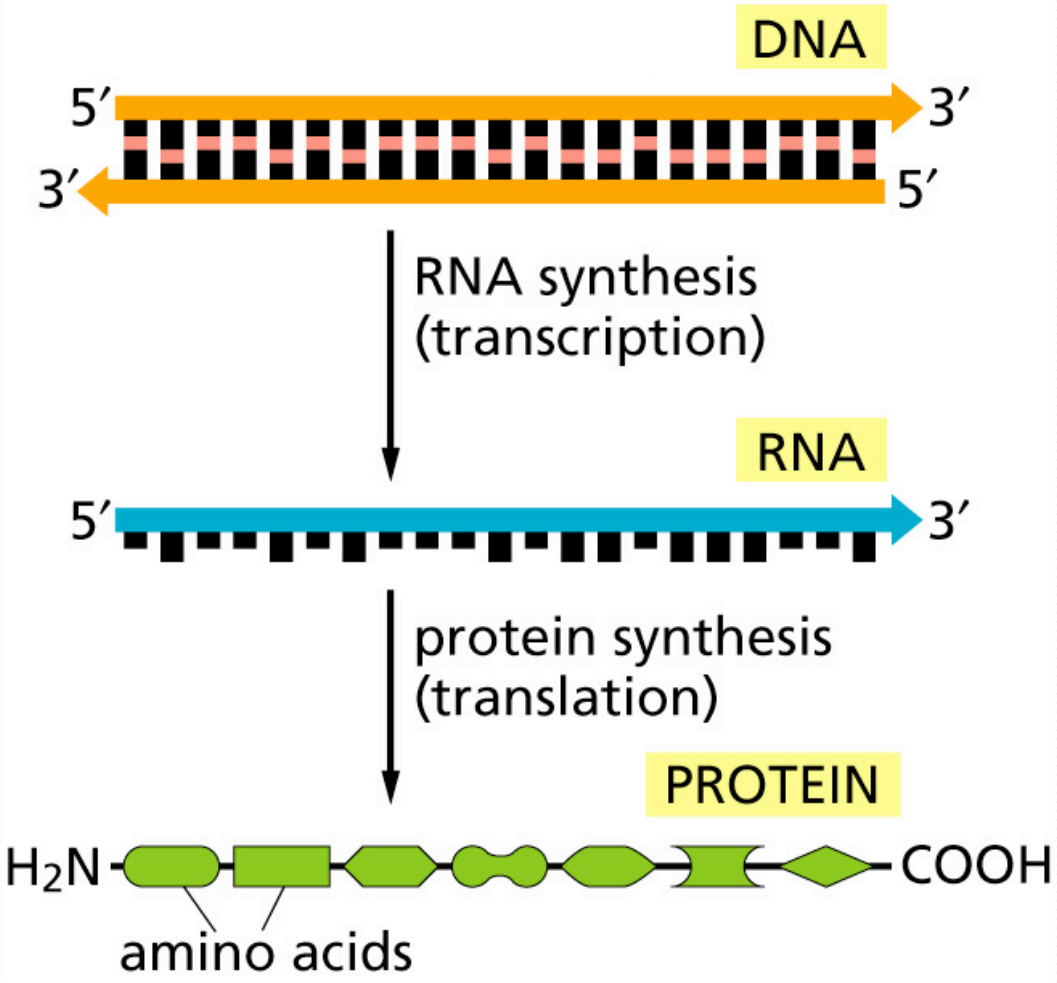
Abstract design

Device Specification

Physical Assembly



BBa_I13521
BBa_I763007
BBa_J3901
BBa_J5526


DNA
5′
3′
3′
5′
RNA synthesis
(transcription)
RNA
5′
3′
protein synthesis
(translation)
PROTEIN
H2N
COOH
amino acids

DNA

- What is the significance of DNA?
  - It is the permanent set of instructions of what proteins to make = **genes**
  - It also contains instructions on WHEN to make the proteins = regulatory regions

- Nucleic acids
  - Polymers (strings) of just 4 nucleotides: the bases are all chemically similar
  - Base pairing between two DNA strands creates a 3 dimensional structure
    - Hydrogen bonding is the primary chemical interaction
  - Base pairing allows for a way to create exact copies

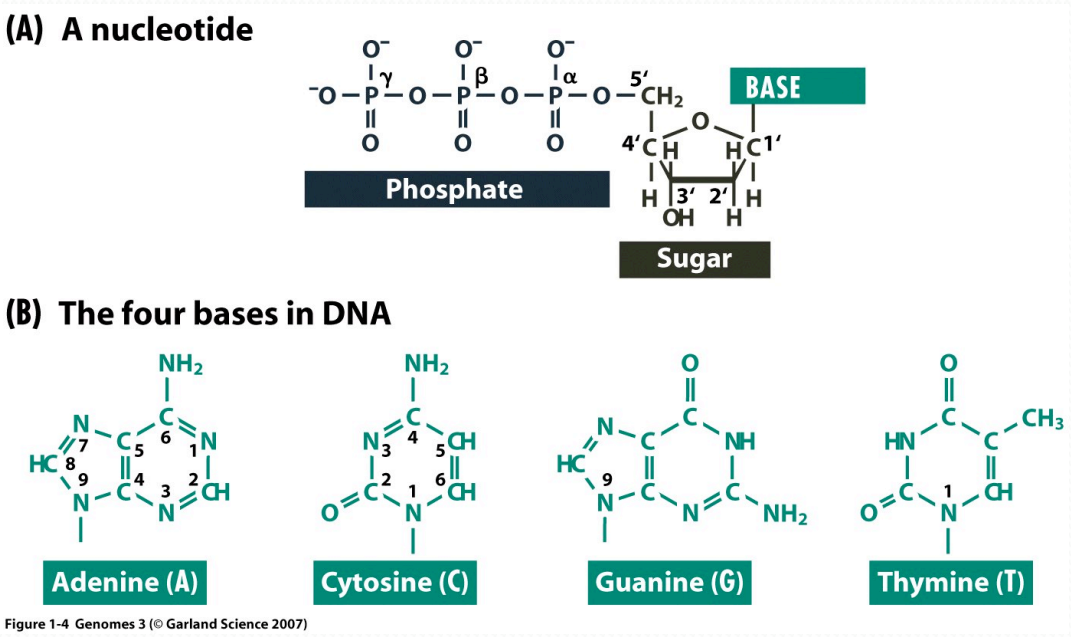


Figure 1-4 Genomes 3 (© Garland Science 2007)

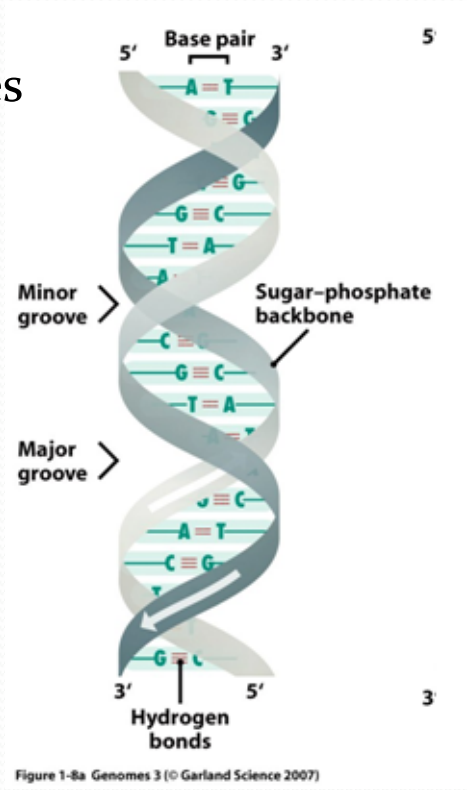
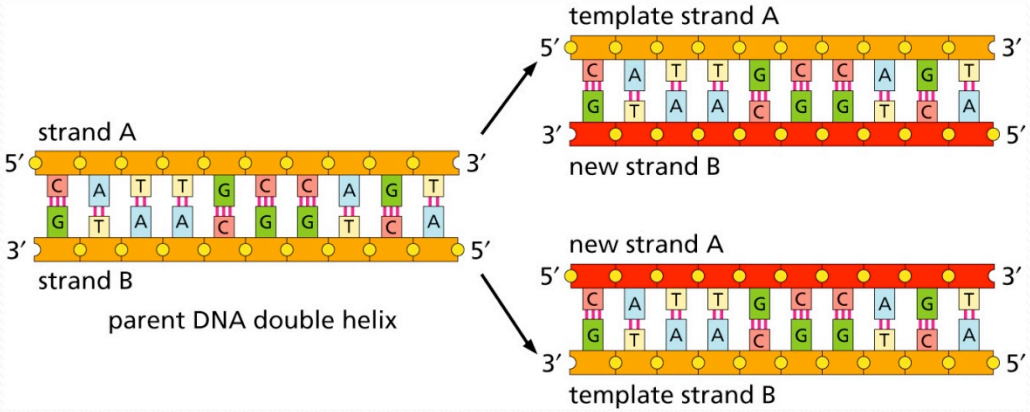


Figure 1-8a Genomes 3 (© Garland Science 2007)

- Nucleic acids are very stable and can produce exact replicas of themselves
  - Double strands separate
  - Each single strand serves as a template to build a complementary copy to produce 2 dsDNA
    - Proteins cannot do this.



DNA is GREAT at carrying information through time

From DNA to protein

- DNA serves as a way to store information through time, through generations
  - This could include information about when and how to build proteins
- Protein serves as a way to build all the components of cells, and to run them
  - Including helping DNA replicate itself so that it can be passed down to a new generation of cells
- If DNA stores information about what proteins to make, and the proteins build the cell, what is the process for translating from DNA language (nucleotides) into protein language (amino acids)?
- Now we can think about the role of RNA
  - RNA allows the instructions to be read and the proteins to be built
  - It translates the language of DNA (nucleotides) into the language of proteins (amino acids)

RNA

- The cell needs a process to
  - Read the instructions
  - Translate the encoded DNA instructions into the building blocks of proteins
  - Build the proteins

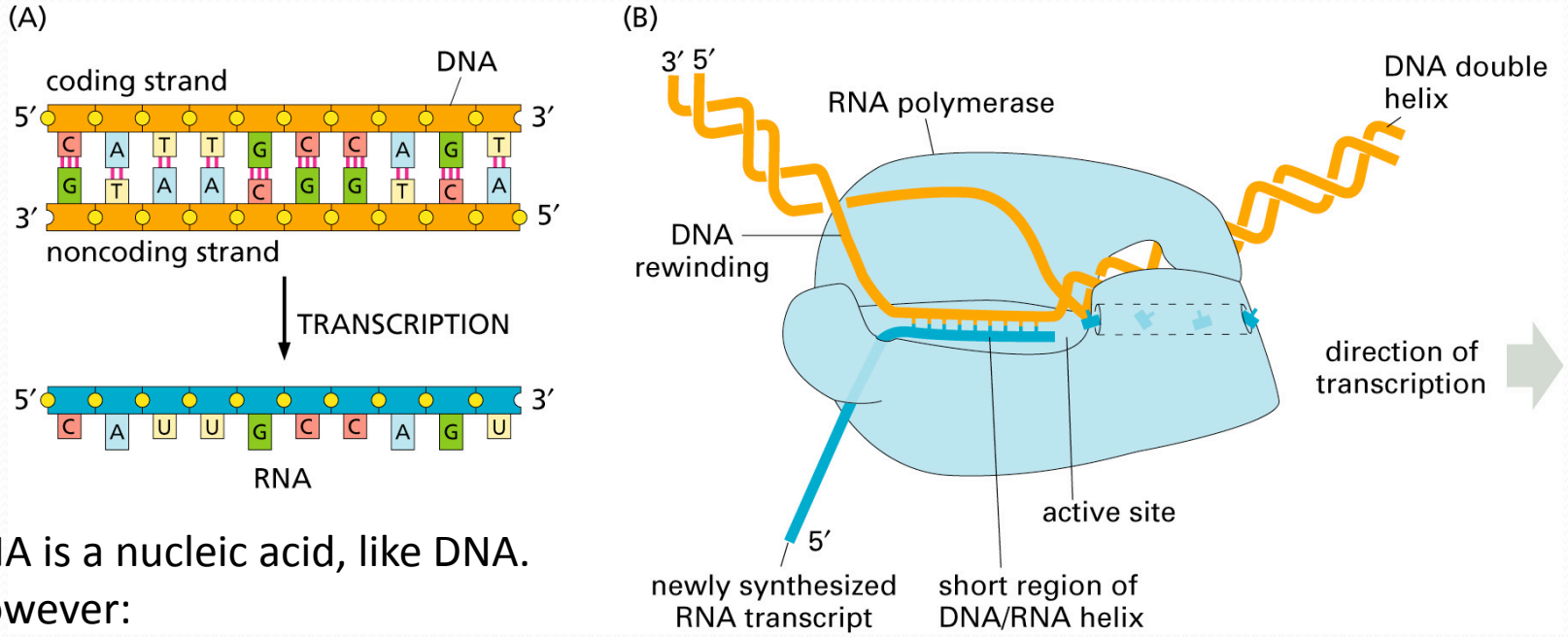
- RNA does all of this
  - Read instructions
  - Translate DNA to AA
  - Build proteins

mRNA – messenger RNA

tRNA – transfer RNA

rRNA – ribosomal RNA

- Messenger RNA
  - A copy of the gene sequence that is mobile and can be carried to the site of protein synthesis



RNA is a nucleic acid, like DNA.

However:

DNA = A T C G

RNA = A **U** C G

CODONS are the translations

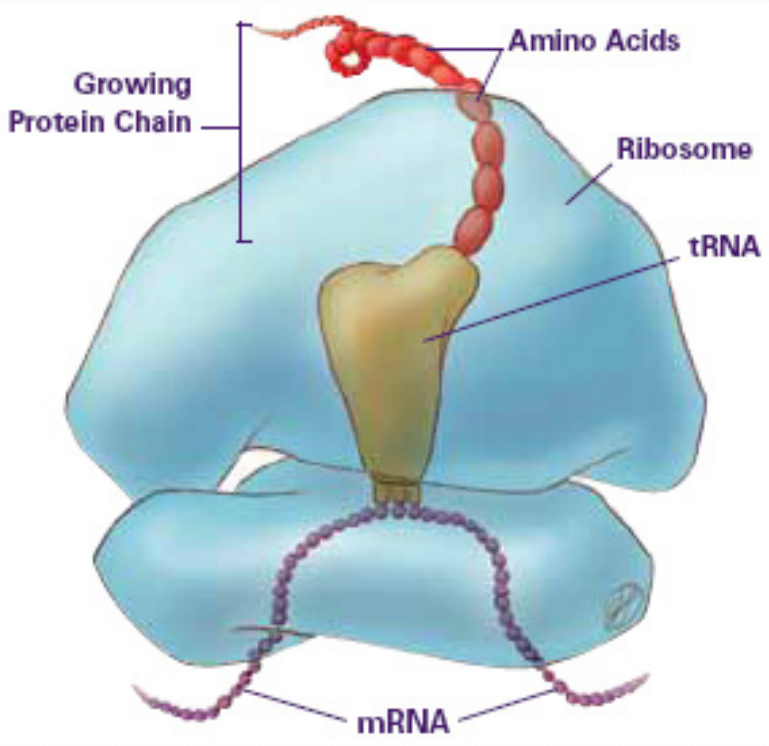
UUU	phe	UCU	ser	UAU	tyr	UGU	cys
UUC		UCC		UAC		UGC	
UUA	leu	UCA		UAA	stop	UGA	stop
UUG		UCG		UAG		UGG	trp
CUU	leu	CCU	pro	CAU	his	CGU	arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	gln	CGA	
CUG		CCG		CAG		CGG	
AUU	ile	ACU	thr	AAU	asn	AGU	ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	lys	AGA	arg
AUG	met	ACG		AAG		AGG	
GUU	val	GCU	ala	GAU	asp	GGU	gly
GUC		GCC		GAC		GGC	
GUA		GCA		GAA	glu	GGA	
GUG		GCG		GAG		GGG	

Figure 1-20 Genomes 3 (© Garland Science 2007)

Each codon represents a tRNA that carries the amino acid shown

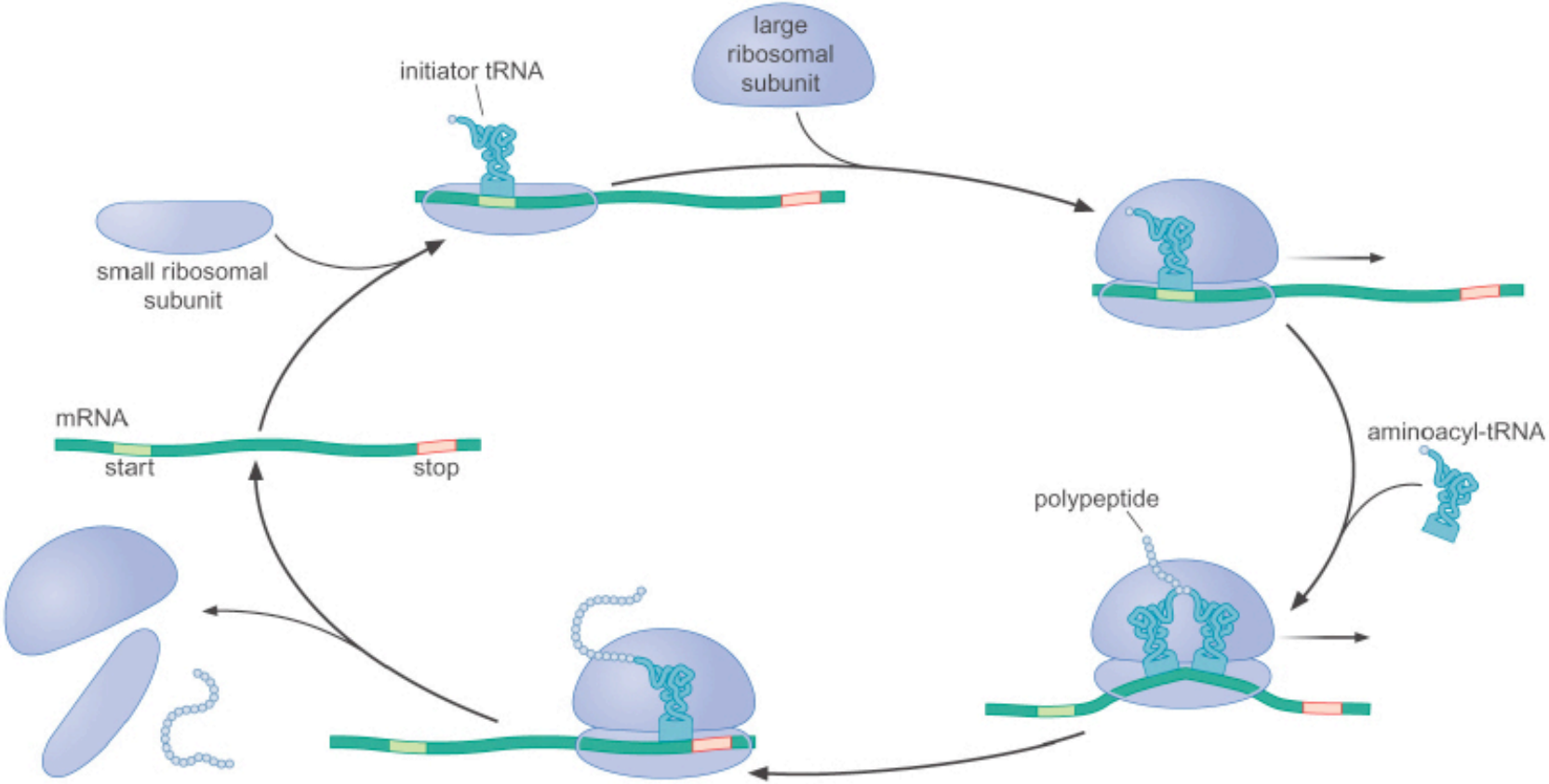
Ribosomal RNA

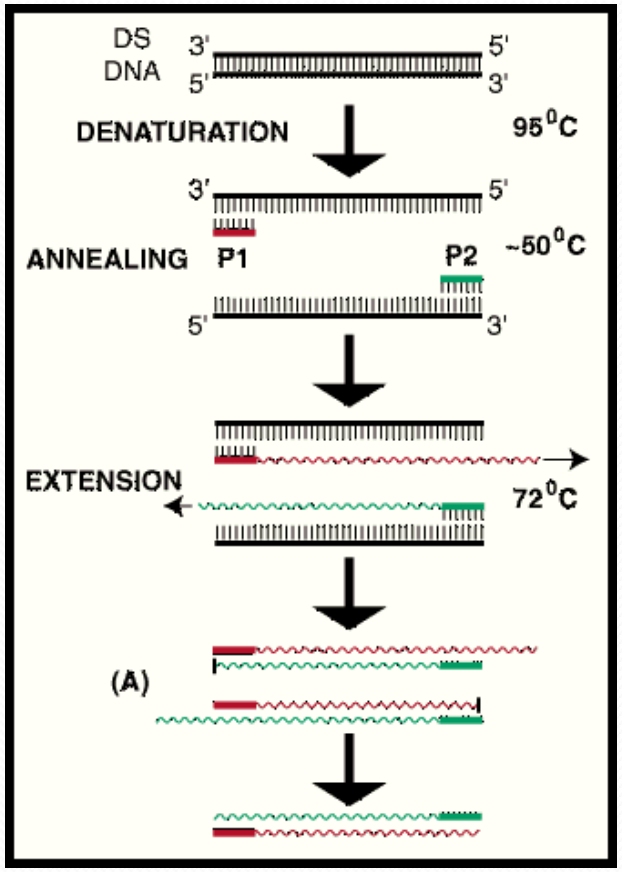
- rRNAs are RNA molecules that join the amino acids together to form a protein
- The ribosome is a piece of machinery made from rRNA and protein



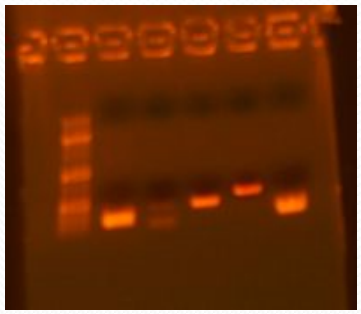
Protein synthesis (translation)

mRNA + tRNA + rRNA + amino acids = protein

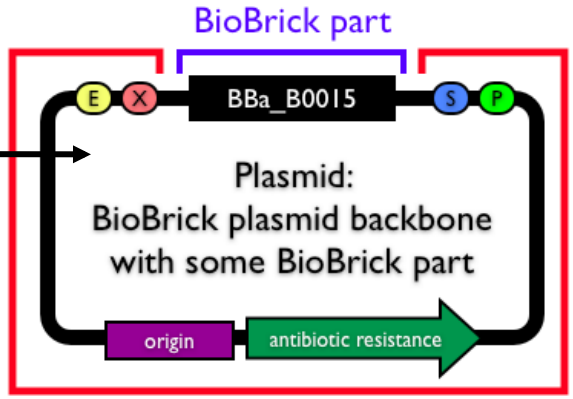




Polymerase Chain Reaction

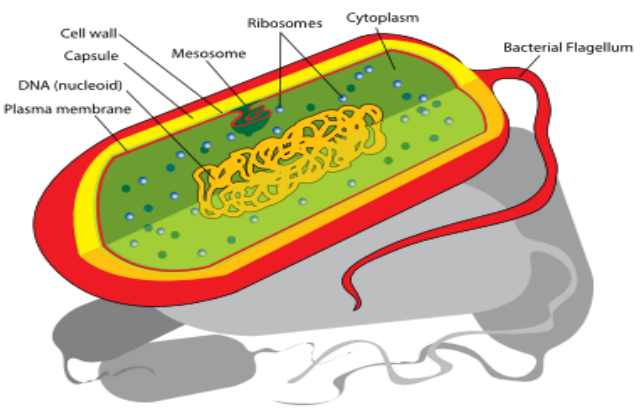


Analytic gel and clean up

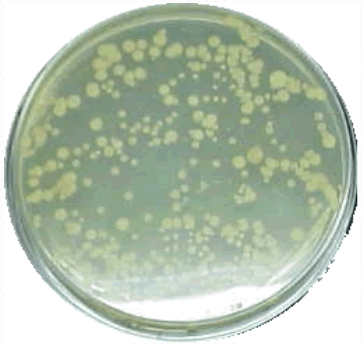


BioBrick plasmid backbone

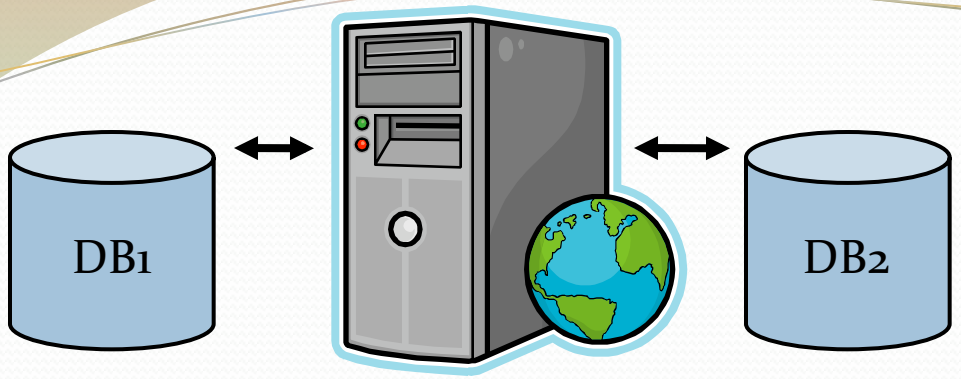
Digest, purify, ligate

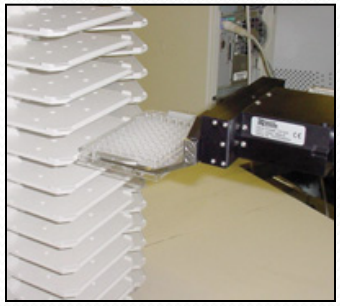


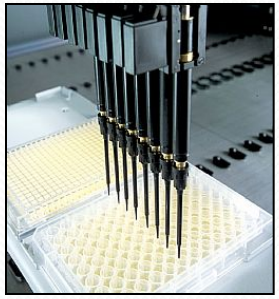
Transform



Plate, pick, and mini-prep





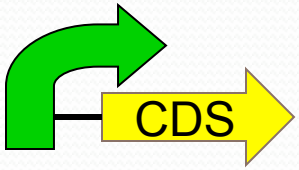


3. Standards
2. Automated Assembly
1. Abstractions

Automatic Sequencing
PCR
Recombinant DNA

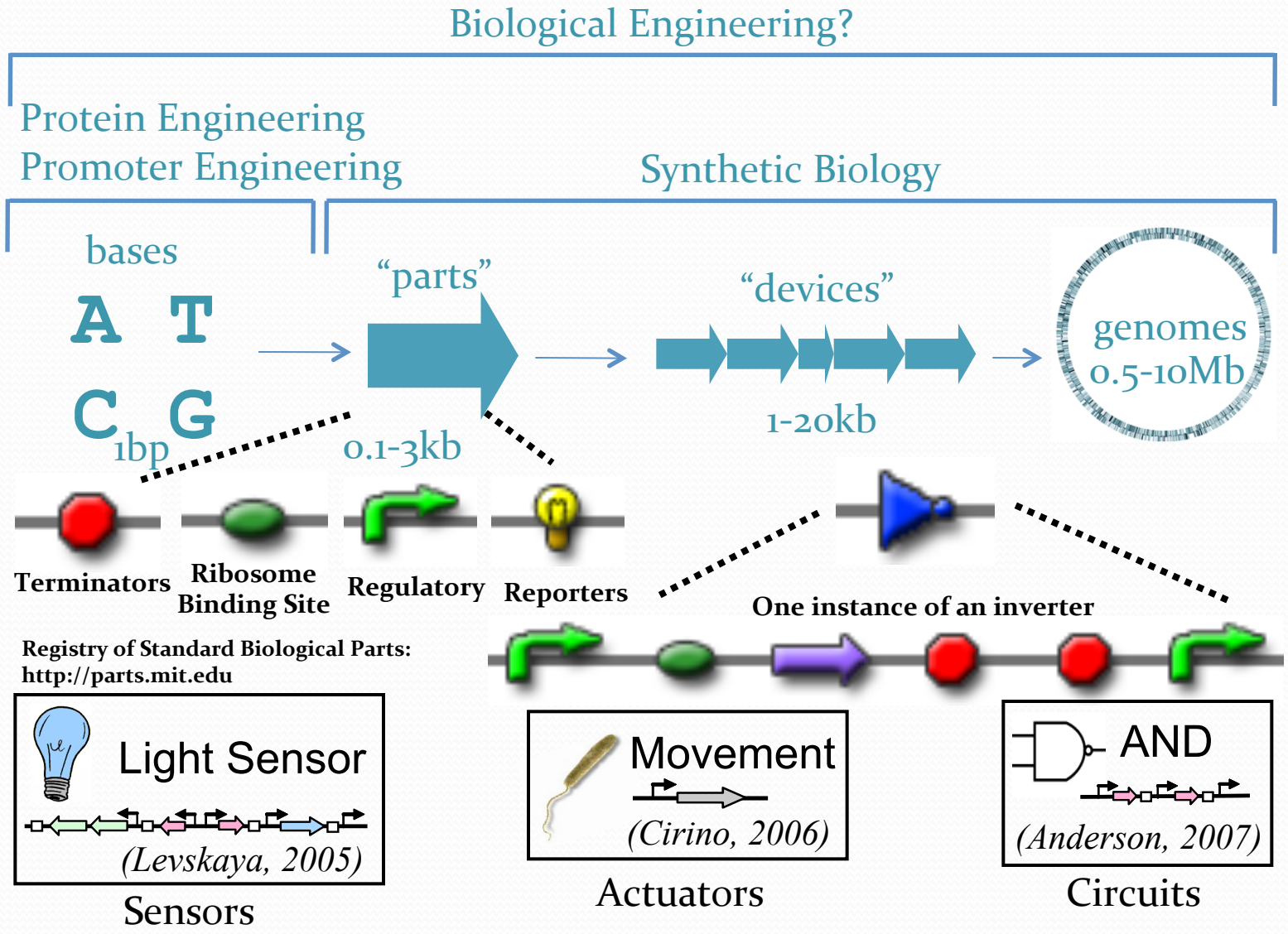
*Drew Endy, Stanford

- DNA Sequence?
  - ATCG
- Function?
- System?

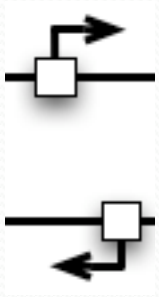


Why This Hard?

Biological uncertainty, meaningful abstractions, consensus on standards


Biological Engineering?
Protein Engineering
Promoter Engineering
Synthetic Biology
bases
A T
C G
1bp
"parts"
0.1-3kb
"devices"
1-20kb
genomes
0.5-10Mb
Terminators
Ribosome Binding Site
Regulatory
Reporters
One instance of an inverter
Registry of Standard Biological Parts:
http://parts.mit.edu
Light Sensor
(Levskaya, 2005)
Sensors
Movement
(Cirino, 2006)
Actuators
AND
(Anderson, 2007)
Circuits

Biological Parts (SBOL Visual Standard)



Promoters

Ribsome Binding Sites

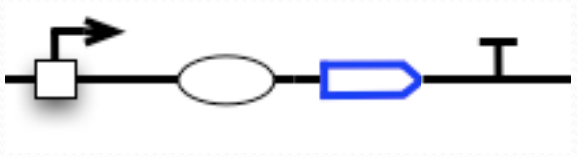
Open Reading Frames

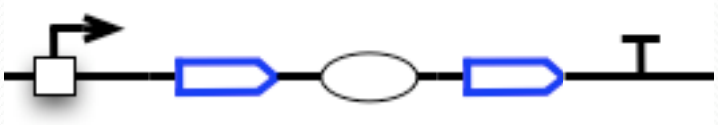
Terminators

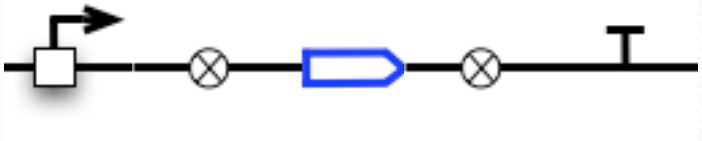
Primers (Oligos)

Restrictions Sites

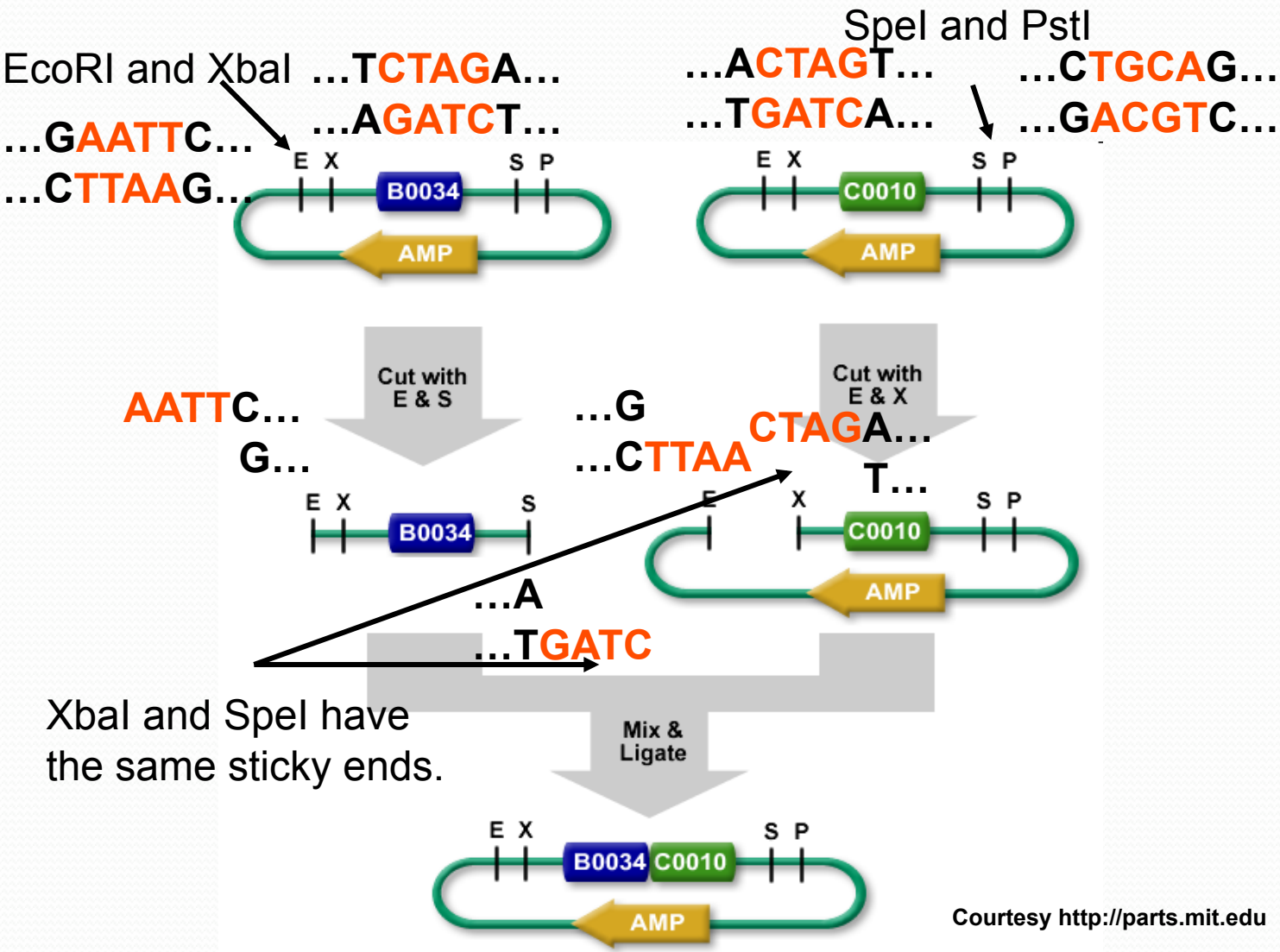
Biological Devices





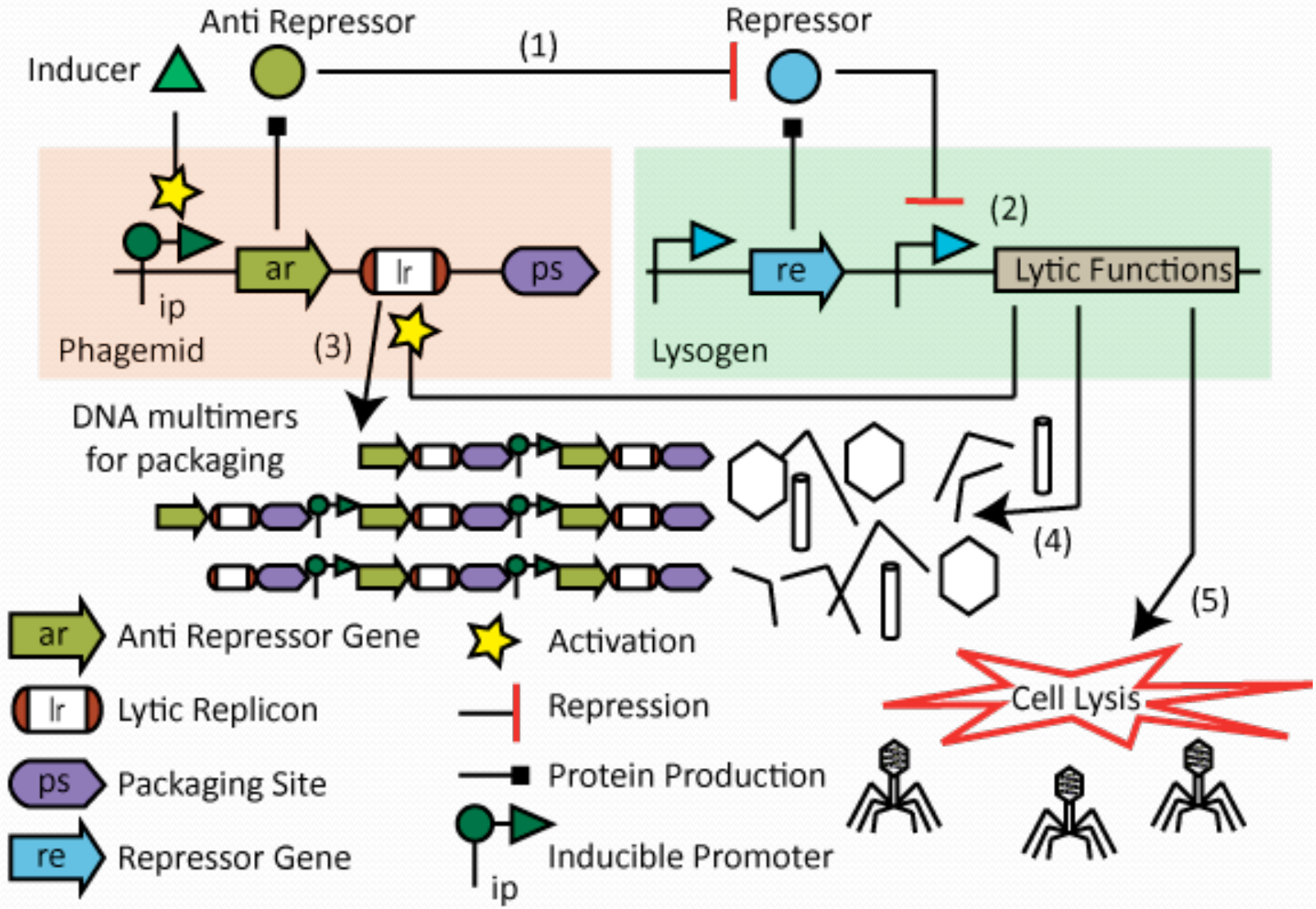


BioBrick Standard Assembly

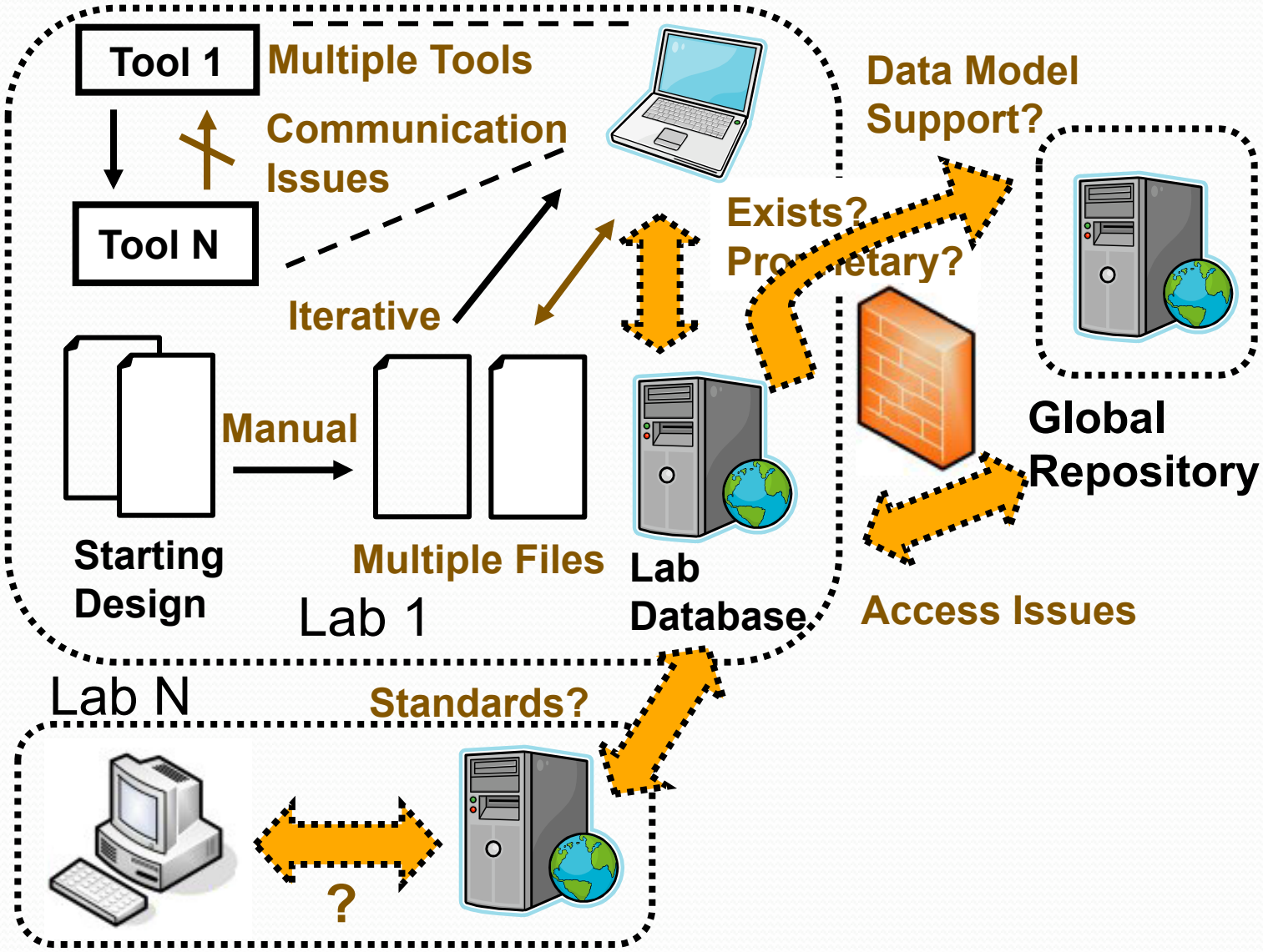


XbaI and SpeI have the same sticky ends.

Courtesy http://parts.mit.edu



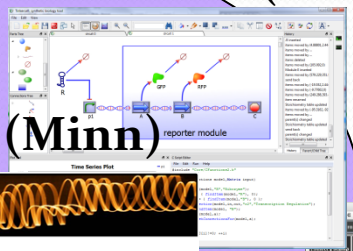
Induction of the phagemid leads to (1) neutralization of the repressor (2) activation of lytic mode (3) amplification of phagemid DNA (4) production of phage particles and (5) cell lysis.


Tool 1
Multiple Tools
Communication Issues
Tool N
Iterative
Manual
Starting Design
Multiple Files
Lab 1
Lab Database
Data Model Support?
Exists?
Pro etary?
Global Repository
Access Issues
Lab N
Standards?
?

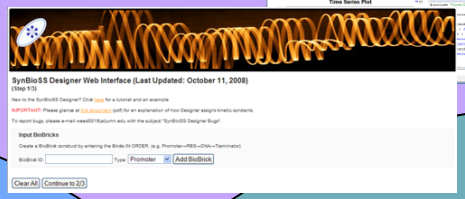
Synthetic Biology Tool Landscape

Computation/Simulation Tools

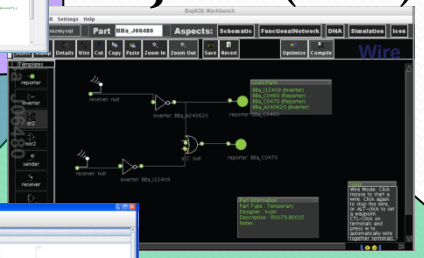
Tinkercell (UW)



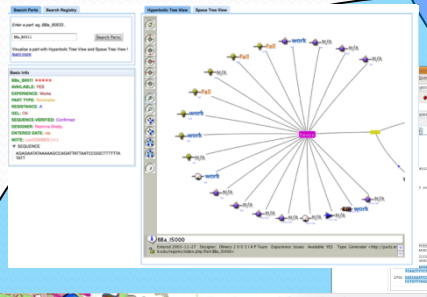
SynBioSS (Minn)



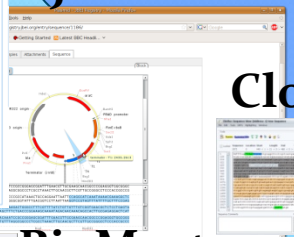
BioJADE (UCB)



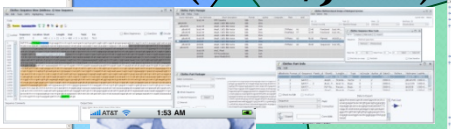
Viz-a-Brick (Davidson-Missouri Western)



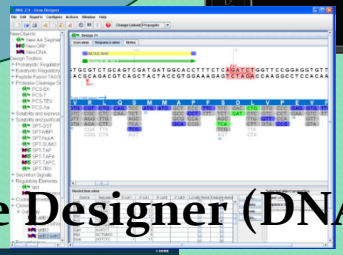
JBEIR



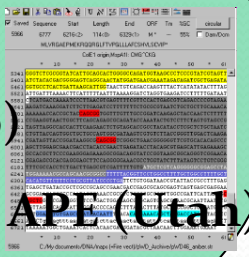
Clotho (UCB/JBEI)



Gene Designer (DNA 2.0)



APE (Utah)

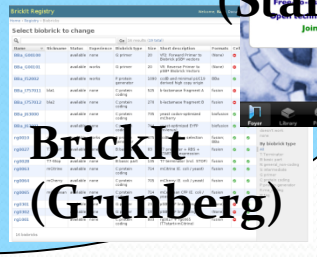


BioMortar (Waterloo)

BB Studio (Stanford)

Parts Registry (MIT)

BrickIt (Grunberg)



GenoCAD (VTech)

BioStudio (Johns Hopkins)

Data Management Tools

Design/Analysis Tools

Abstract design

Clotho Core API

Revamped Clotho Core and Data Model

Device Specification

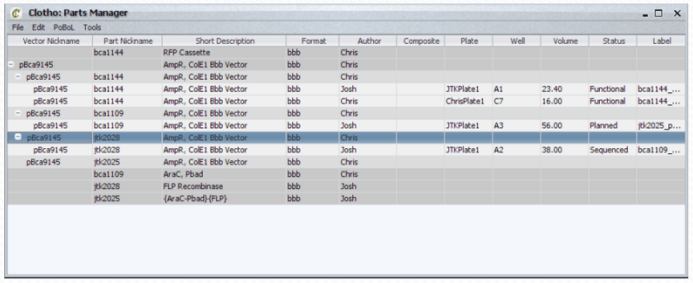
Eugene

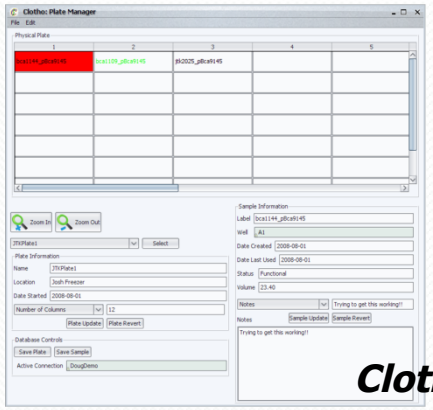
Physical Assembly

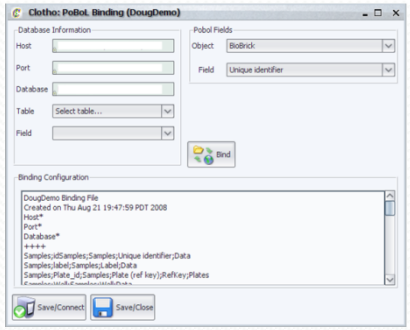
Kepler

Clotho – Data Management/Design

***Clotho* Parts Manager**

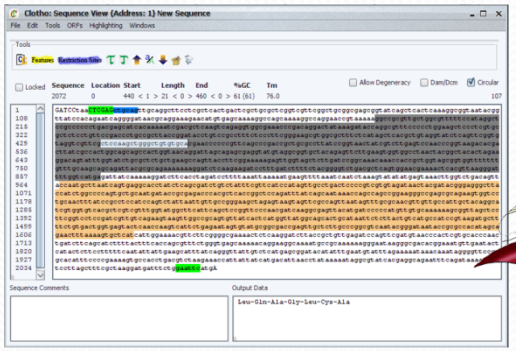


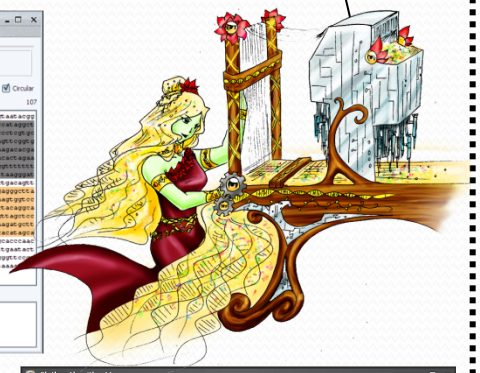




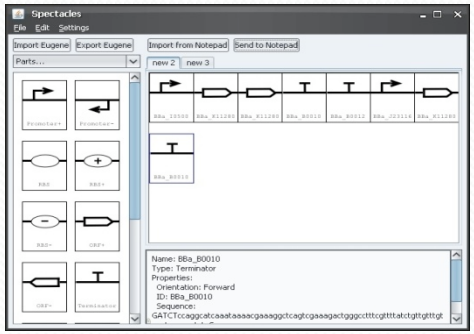
***Clotho* Binding Manager**

***Clotho* Plate Manager**

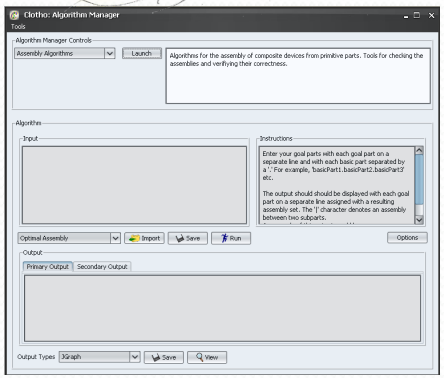




***Clotho* Sequence View**



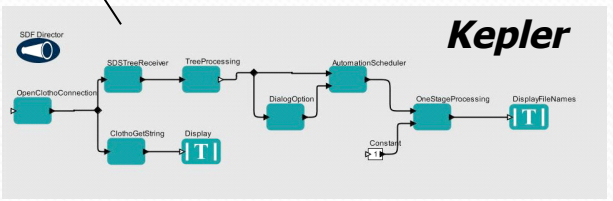
Spectacles



***Clotho* Algorithm Manager**

Pros: PlugIn Based Tools, Flexible Data Retrieval, Automatic Assembly, Part/Device Design

Eugene

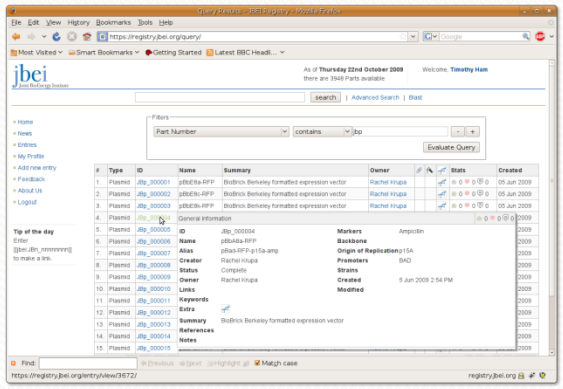


Cons: Still in development, requires data source and internet connection, most powerful with Part/Device methodology

http://www.clothocad.org

Design Flow

Data API



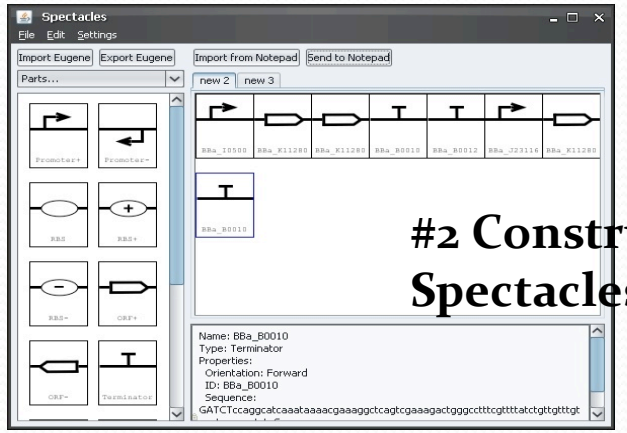
Eugene

#1 Specification

Comm. API

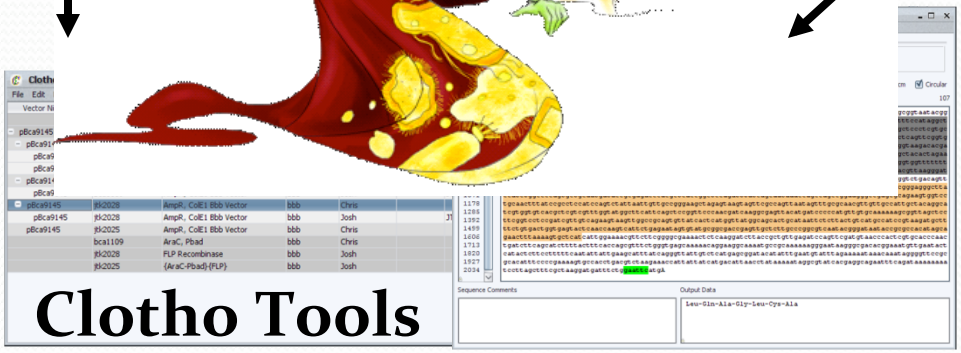
Data API

Data API

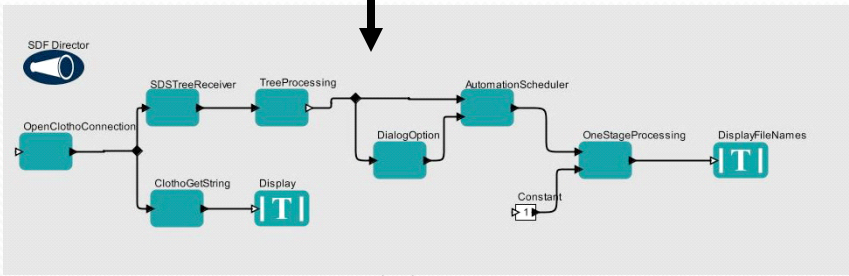


#2 Construction
Spectacles

Comm. API

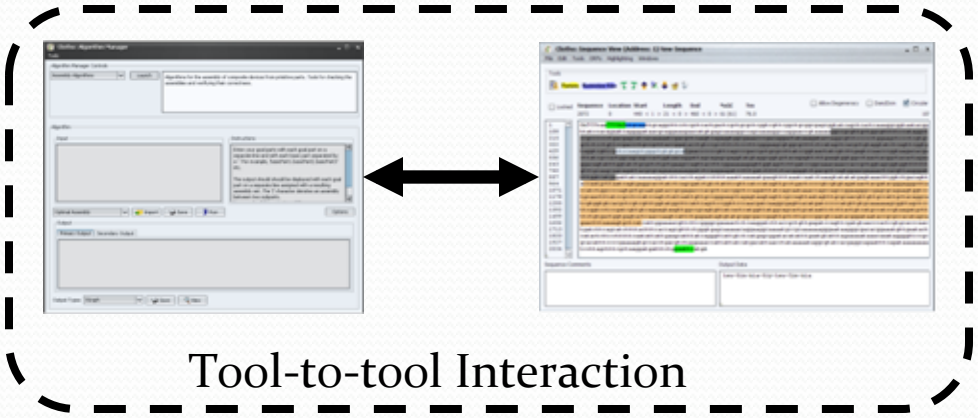


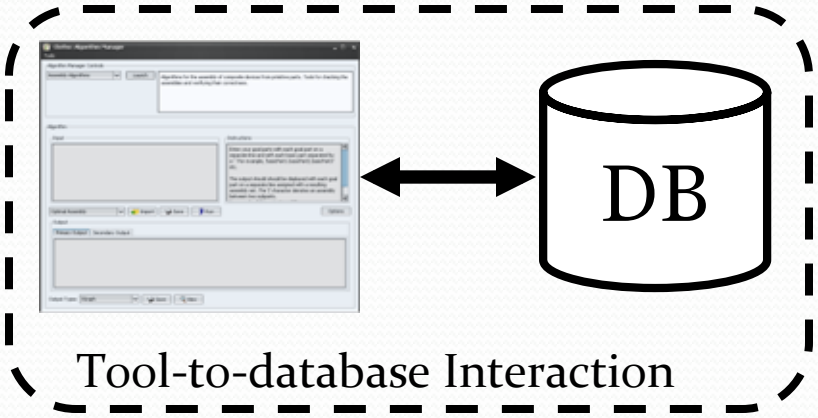
Clotho Tools

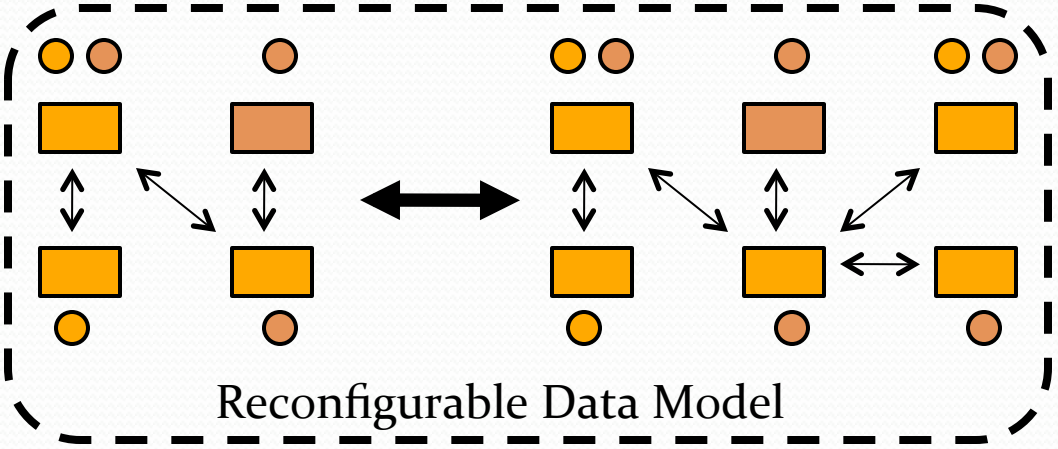


#3 Assembly
Kepler Workflow

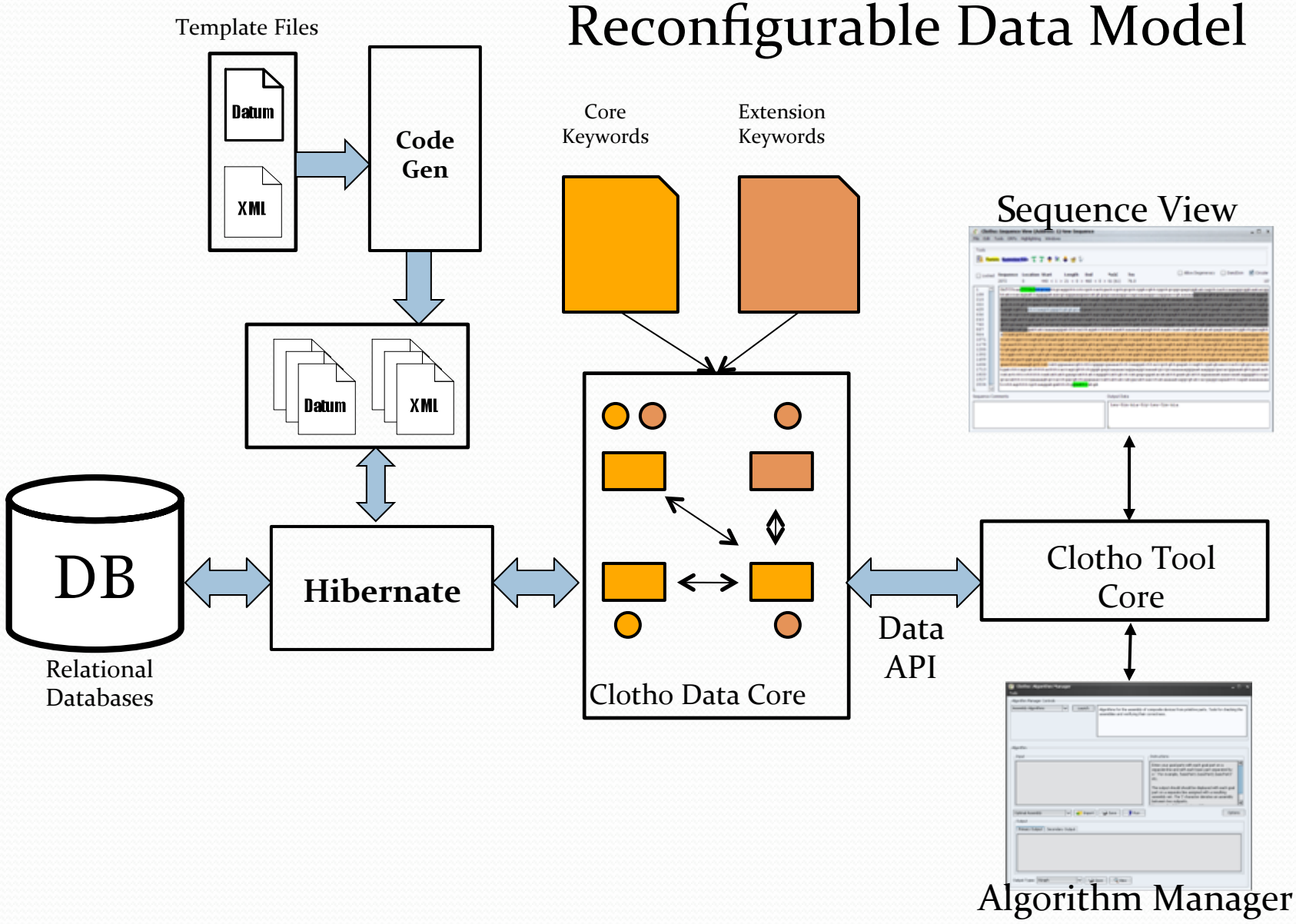
Clotho Infrastructure





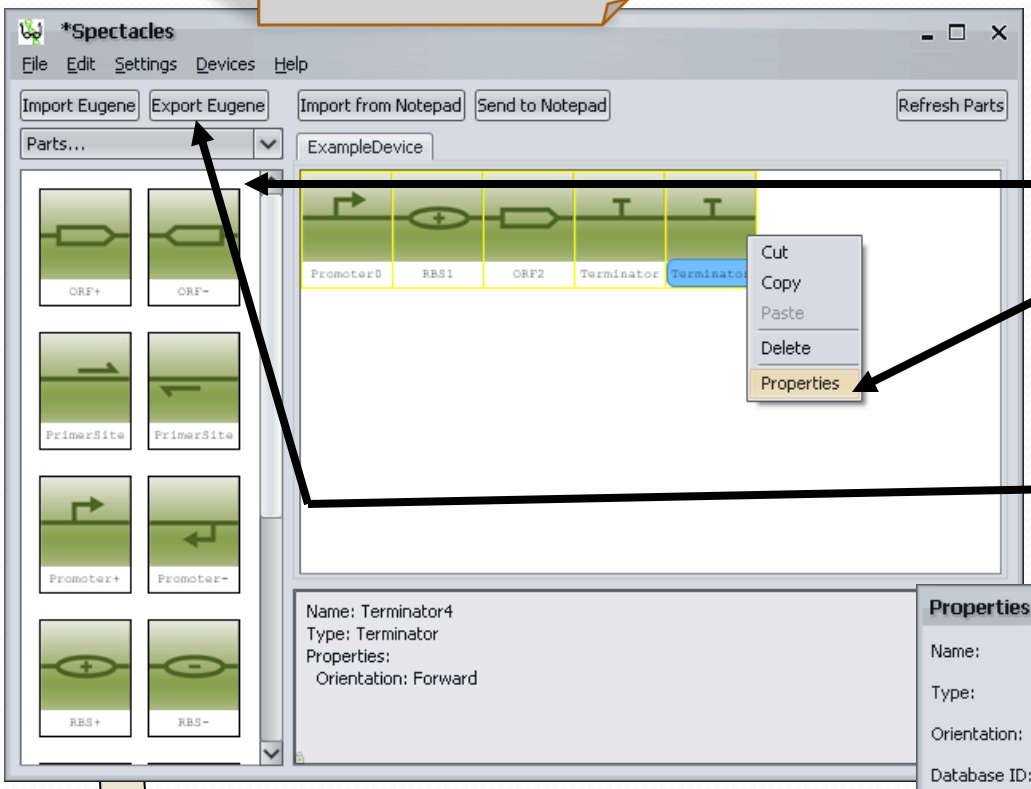


Reconfigurable Data Model



Abstract Design

Eugene

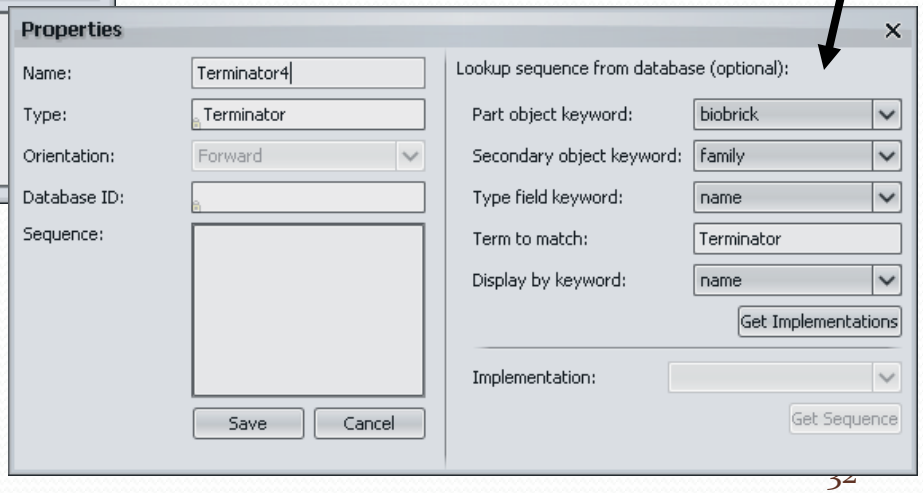


Choose a part

Modify name, sequence, or implementation

Map to physical part in database

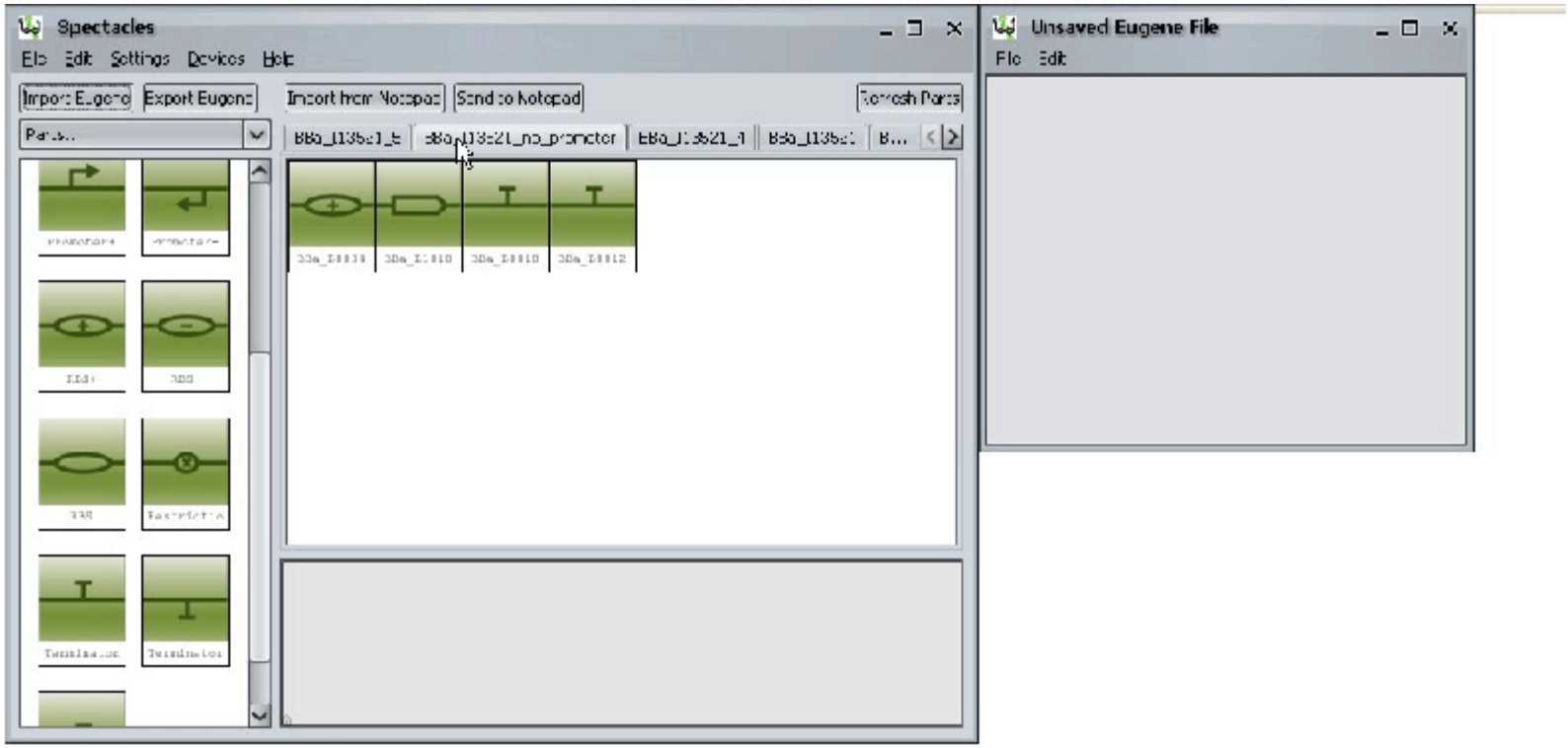
Export to Eugene



Tool API

Clotho
- Sequence View
- Algorithm Manager

Spectacles Demo



Device Specification

Realize an abstract design from Spectacles:

<part declarations> → Specify instances of physical parts

<rule declarations> → Specify rules on how parts can interact with each other on devices

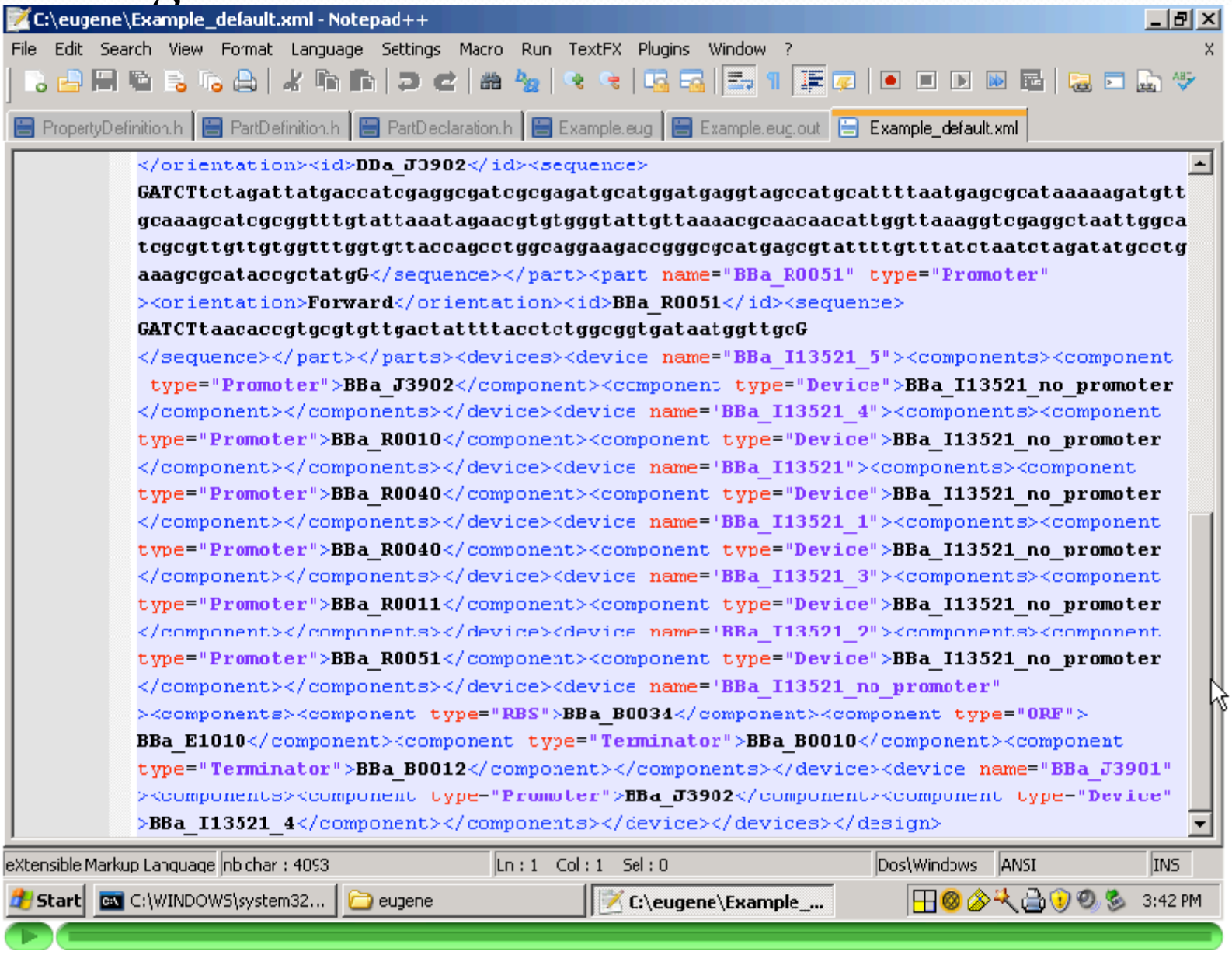
<device declarations> → Construct devices while enforcing rules

<function calls on devices> → Perform functions on devices

Eugene Compiler → XML → Tool X

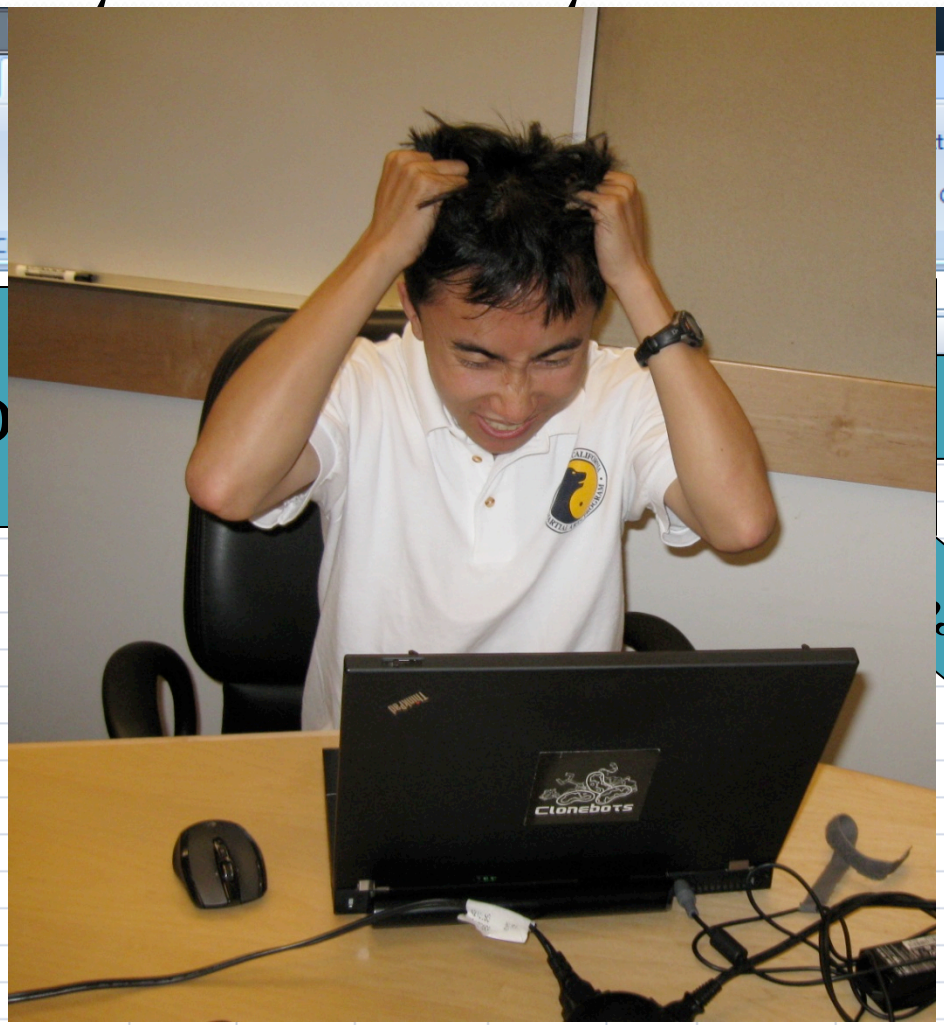
Eugene Header Files → XML To Eugene Header Tool → XML → Database Of Parts

Eugene Demo

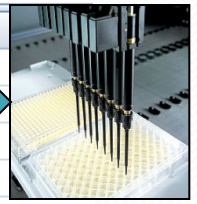

C:\eugene\Example_default.xml - Notepad++
File Edit Search View Format Language Settings Macro Run TextFX Plugins Window ?
PropertyDefinition.h
PartDefinition.h
PartDeclaration.h
Example.eug
Example.eug.out
Example_default.xml
</orientation><id>BBa_J3902</id><sequence>
GATCTtctagattatgaccatcgaggcgatcgcgagatgcatggatgaggtagccatgcattttaatgagcgcataaaaagatgtt
gcaaagcatcgcggtttgtattaaatagaacgtgtgggtattgttaaaacgcaacaacattggttaaaggtcgaggctaattggca
tcgcgttgttgtggtttggtgttaccagcctggcaggaagaccgggcgcatgagcgtattttgtttatctaatctagatatgcctg
aaagcgcataccgctatgG</sequence></part><part name="BBa_R0051" type="Promoter"
><orientation>Forward</orientation><id>BBa_R0051</id><sequence>
GATCTtaacaccgtgcgtgttgactattttacctctggcggtgataatggttgcG
</sequence></part></parts><devices><device name="BBa_I13521_5"><components><component
type="Promoter">BBa_J3902</component><component type="Device">BBa_I13521_no_promoter
</component></components></device><device name="BBa_I13521_4"><components><component
type="Promoter">BBa_R0010</component><component type="Device">BBa_I13521_no_promoter
</component></components></device><device name="BBa_I13521"><components><component
type="Promoter">BBa_R0040</component><component type="Device">BBa_I13521_no_promoter
</component></components></device><device name="BBa_I13521_1"><components><component
type="Promoter">BBa_R0040</component><component type="Device">BBa_I13521_no_promoter
</component></components></device><device name="BBa_I13521_3"><components><component
type="Promoter">BBa_R0011</component><component type="Device">BBa_I13521_no_promoter
type="Promoter">BBa_R0051</component><component type="Device">BBa_I13521_no_promoter
</component></components></device><device name="BBa_I13521_no_promoter"
><components><component type="RBS">BBa_B0034</component><component type="ORF">
BBa_E1010</component><component type="Terminator">BBa_B0010</component><component
type="Terminator">BBa_B0012</component></components></device><device name="BBa_J3901"
>BBa_I13521_4</component></components></device></devices></design>
eXtensible Markup Language nb char : 4093 Ln : 1 Col : 1 Sel : 0 Dos\Windows ANSI INS
Start C:\WINDOWS\system32... eugene C:\eugene\Example_... 3:42 PM

Physical Assembly

Clotho



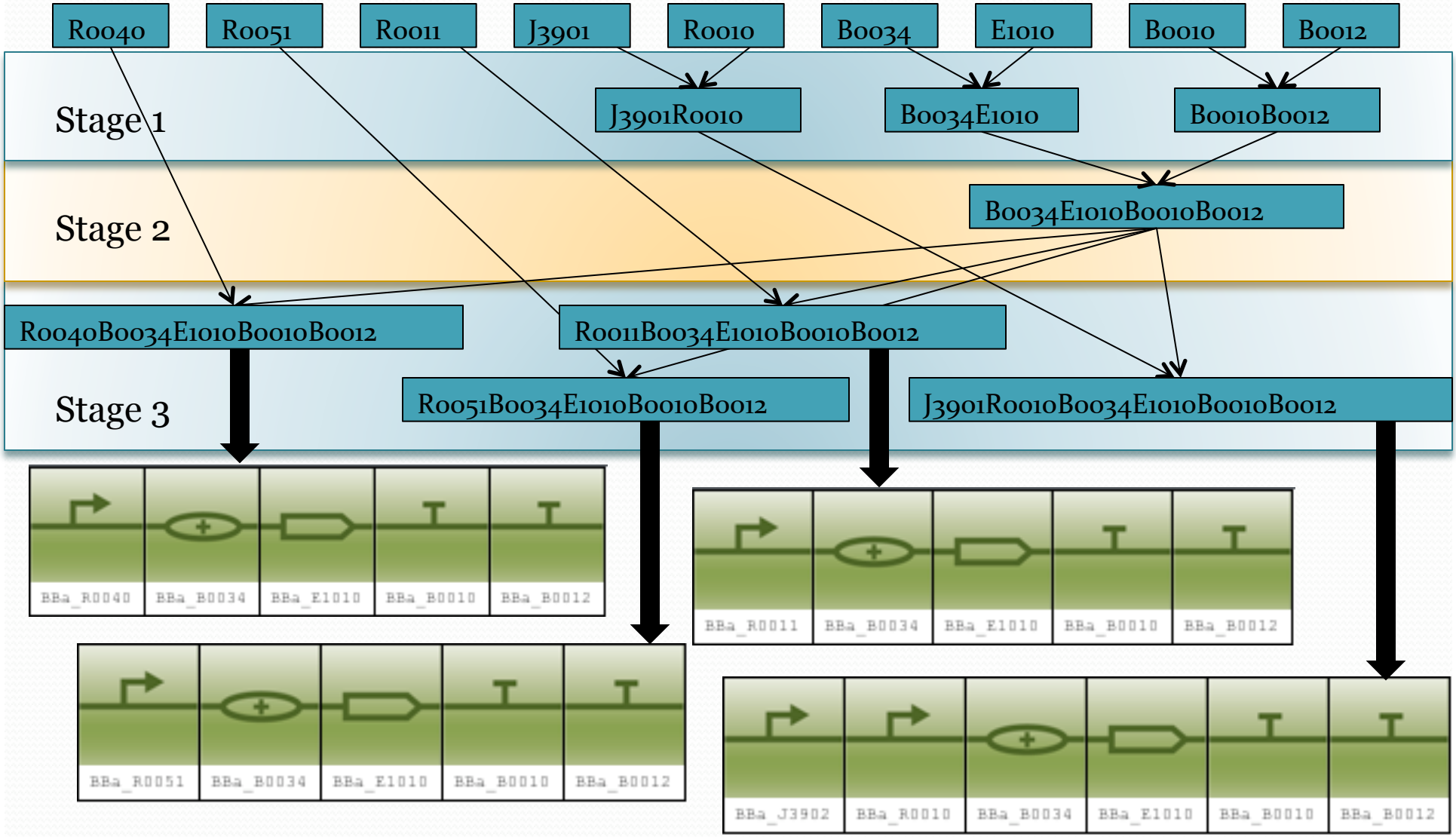
Files for robot



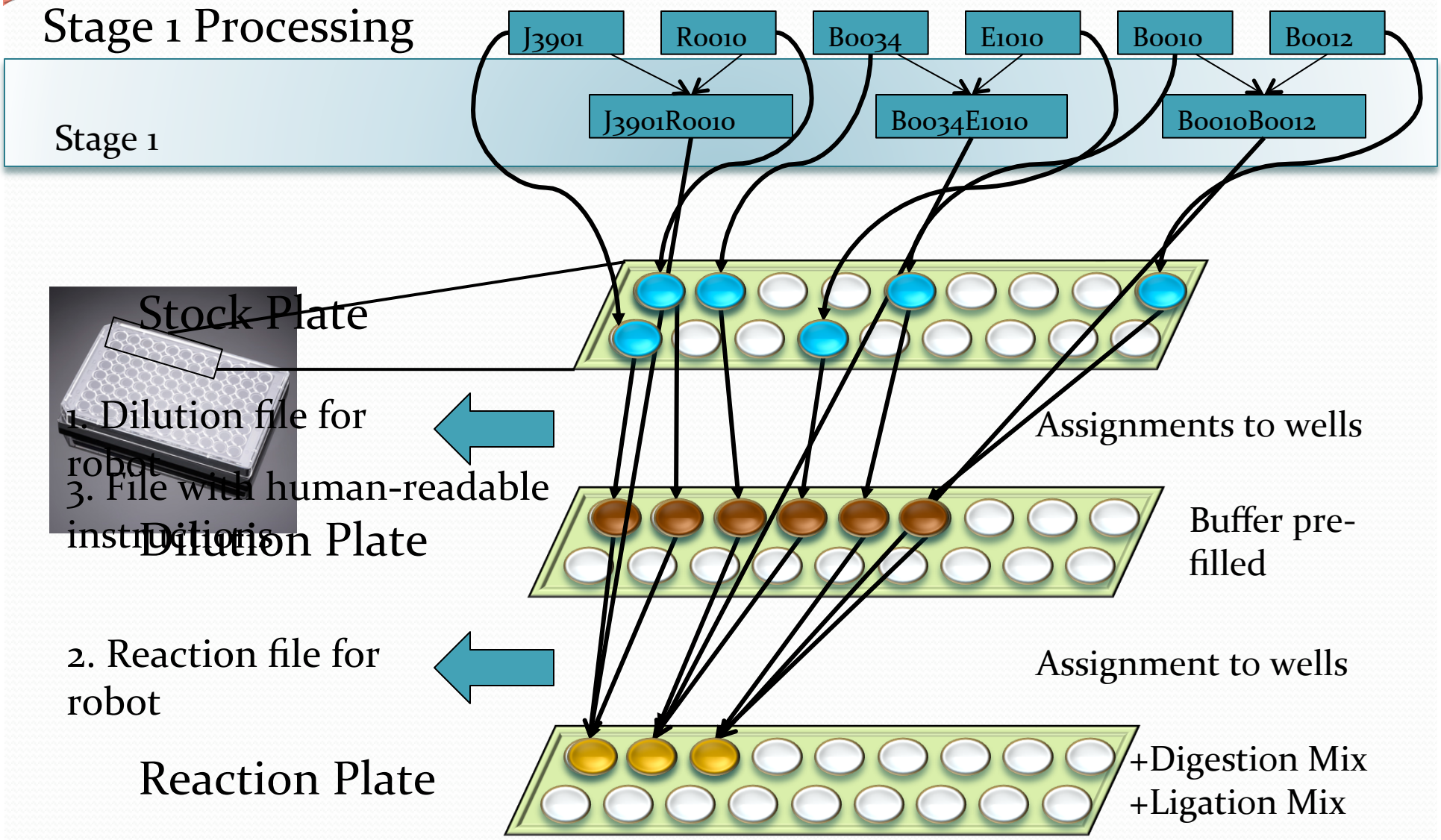
an Instructions

6	P2								0	0
7	P2								0	0
8	P2								0	0
9	P2								0	0
10	P2								0	0
11	P2								0	0
12	P2								0	0
13	P2								0	0
14	P2								0	0
15	P2								0	0
16	P2								0	0
17	P2								0	0
18	P2								0	0
19	P2	C12	B5	P3	A6	A5	P4	B6	0	0
20	P2	C12	B7	P3	A6	A5	P4	B7	0	0
21	P2	C12	B11	P3	A6	A5	P4	B8	0	0
22	P2	C12	B2	P3	A6	A5	P4	B9	0	0
23	P2	D1	B4	P3	A6	A5	P4	B10	0	0

Assembly Graph



Stage 1 Processing



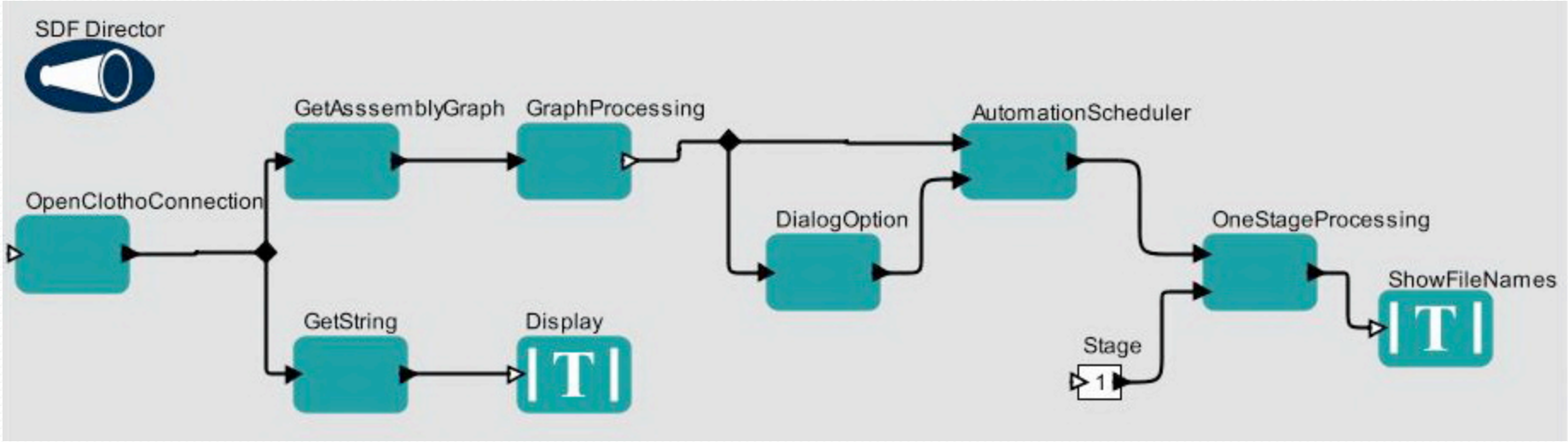
Kepler

Workflow design environment

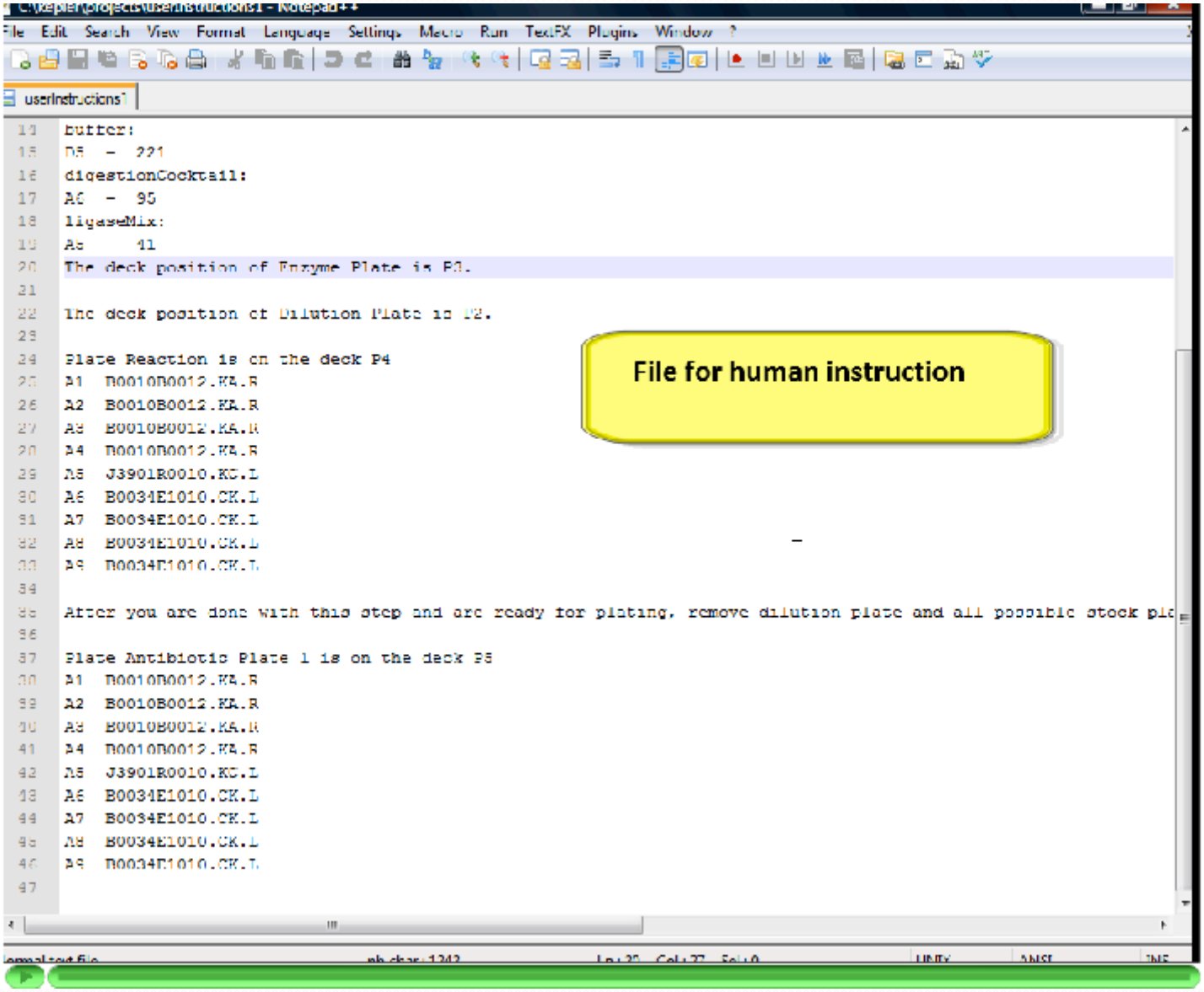
- Visual
- Extensible

Director: controls the execution of actors

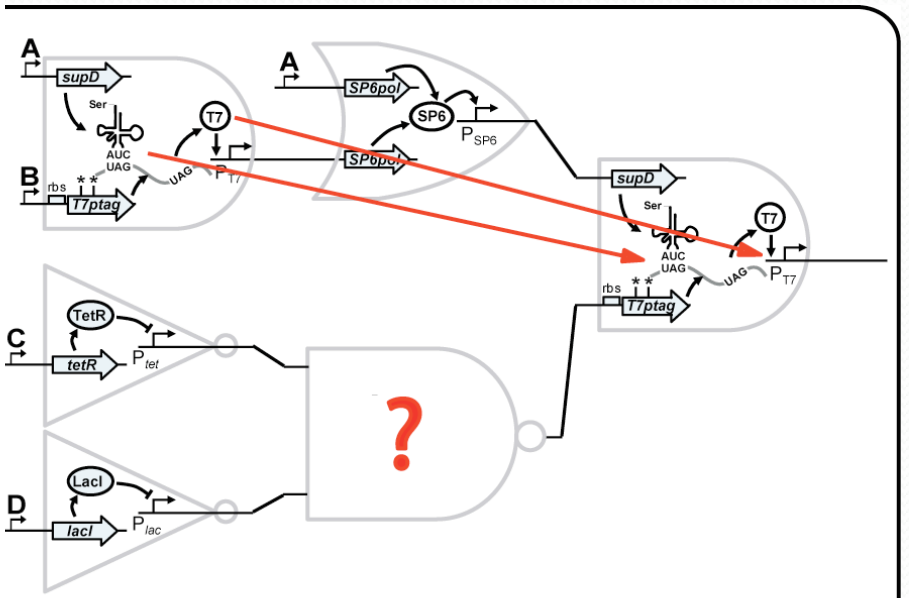
Actor: a process step in the workflow



Kepler Demo


File Edit Search View Format Language Settings Macro Run TextFX Plugins Window ?
userInstructions*
butter:
D5 - 221
digestionCocktail:
A6 - 95
ligaseMix:
A5 41
The deck position of Enzyme Plate is P1.
The deck position of Dilution Plate is P2.
Plate Reaction is on the deck P4
A1 B0010B0012.KA.R
A2 B0010B0012.KA.R
A3 B0010B0012.KA.R
A4 B0010B0012.KA.R
A5 J3901R0010.KC.L
A6 B0034E1010.CK.L
A7 B0034E1010.CK.L
A8 B0034E1010.CK.L
After you are done with this step and are ready for plating, remove dilution plate and all possible stock pla
Plate Antibiotic Plate 1 is on the deck P5
A1 B0010B0012.KA.R
A2 B0010B0012.KA.R
A3 B0010B0012.KA.R
A4 B0010B0012.KA.R
A5 J3901R0010.KC.L
A6 B0034E1010.CK.L
A7 B0034E1010.CK.L
A8 B0034E1010.CK.L
File for human instruction

Example Design Automation Opportunities



Conclusions

- Synthetic biology is an engineering approach to building biological systems
- Design automation can have a prominent role in the advancement of the field
- The key is to leverage the strengths of design automation while respecting biology's unique properties.
- The field is in its infancy so there is a lot of room to make key contributions.