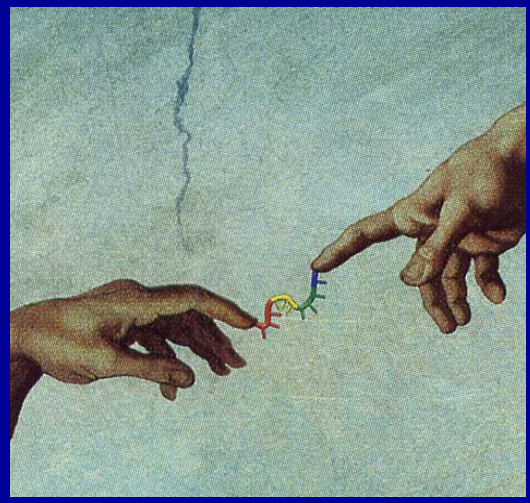
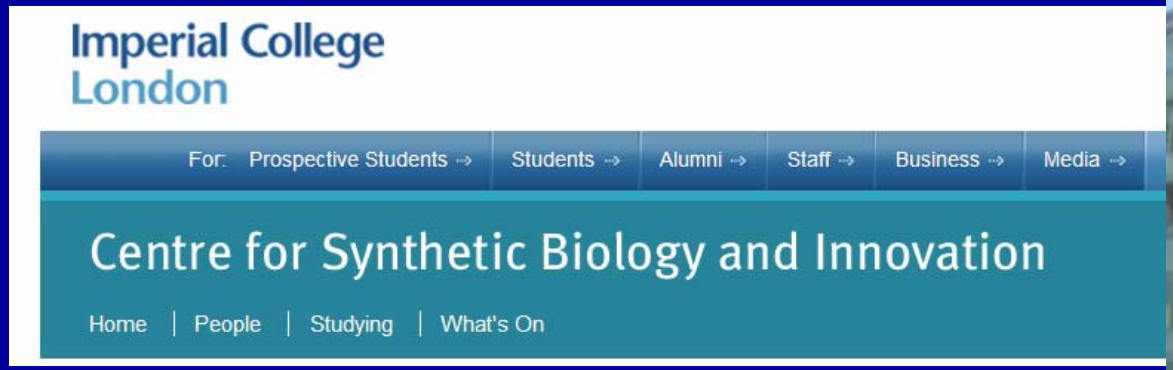
Imperial College London



Systematic Design

Richard I Kitney

Co director - UK National Centre for Synthetic Biology and Innovation



http://www3.imperial.ac.uk/syntheticbiology

Key Points

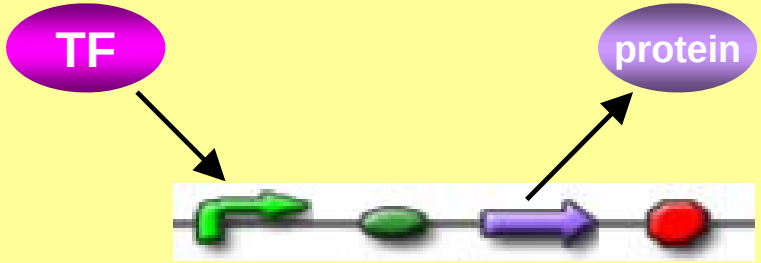
The endpoint of Systems Biology is the analysing biological systems

The endpoint of Synthetic Biology is industrialisation

So how do we achieve this objective

Engineering v Biology

Modularity, Characterisation, Standardisation



Typical gene transcription module

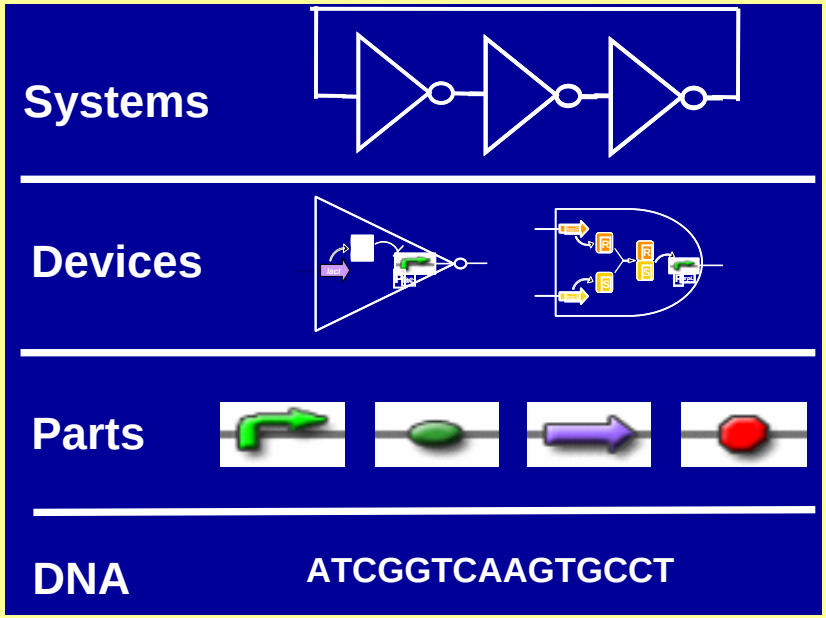

Ribosome binding site


Protein coding sequence


Terminator

Transcription factor

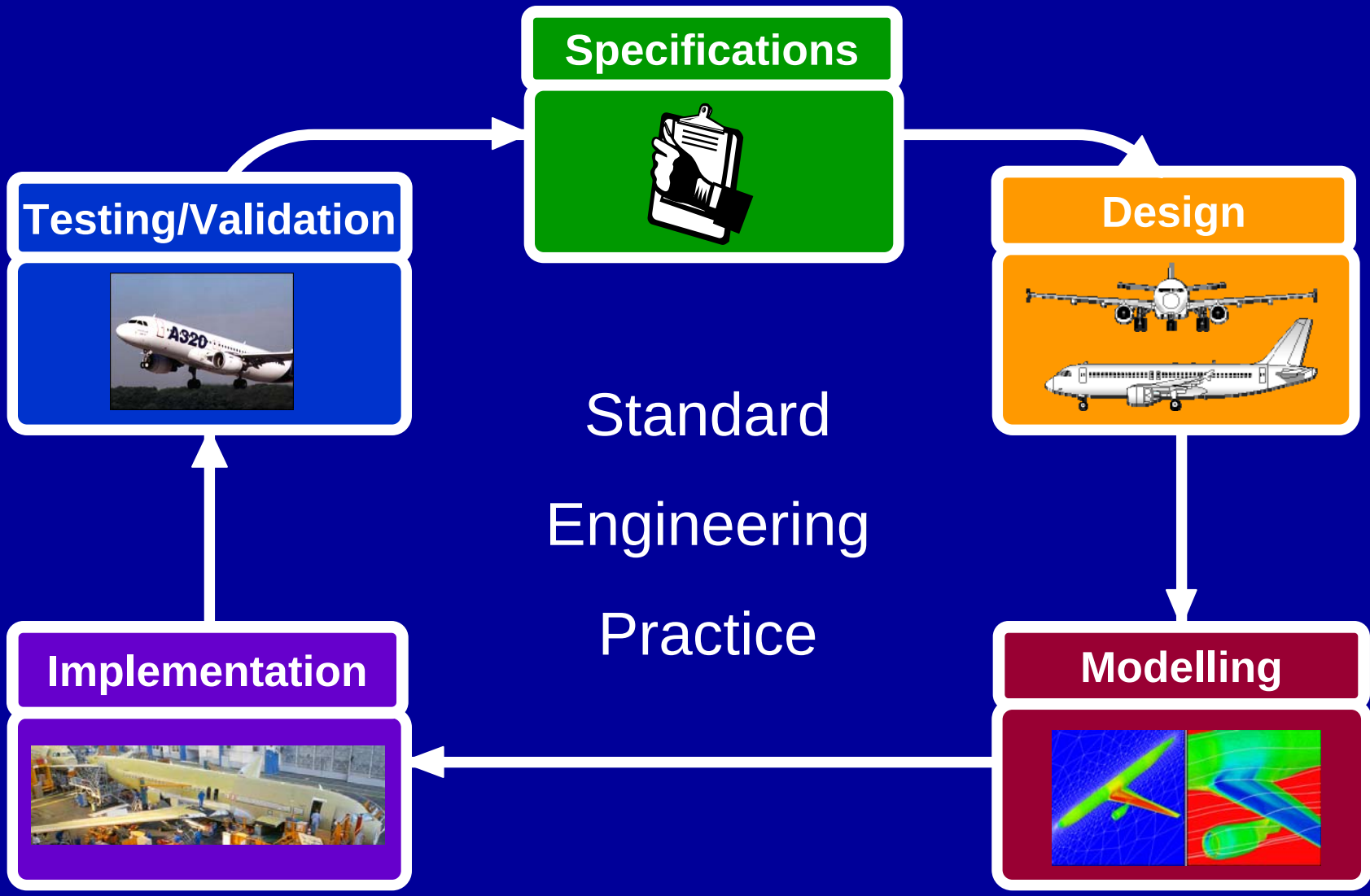
A hierarchy for synthetic biology



Systematic Design

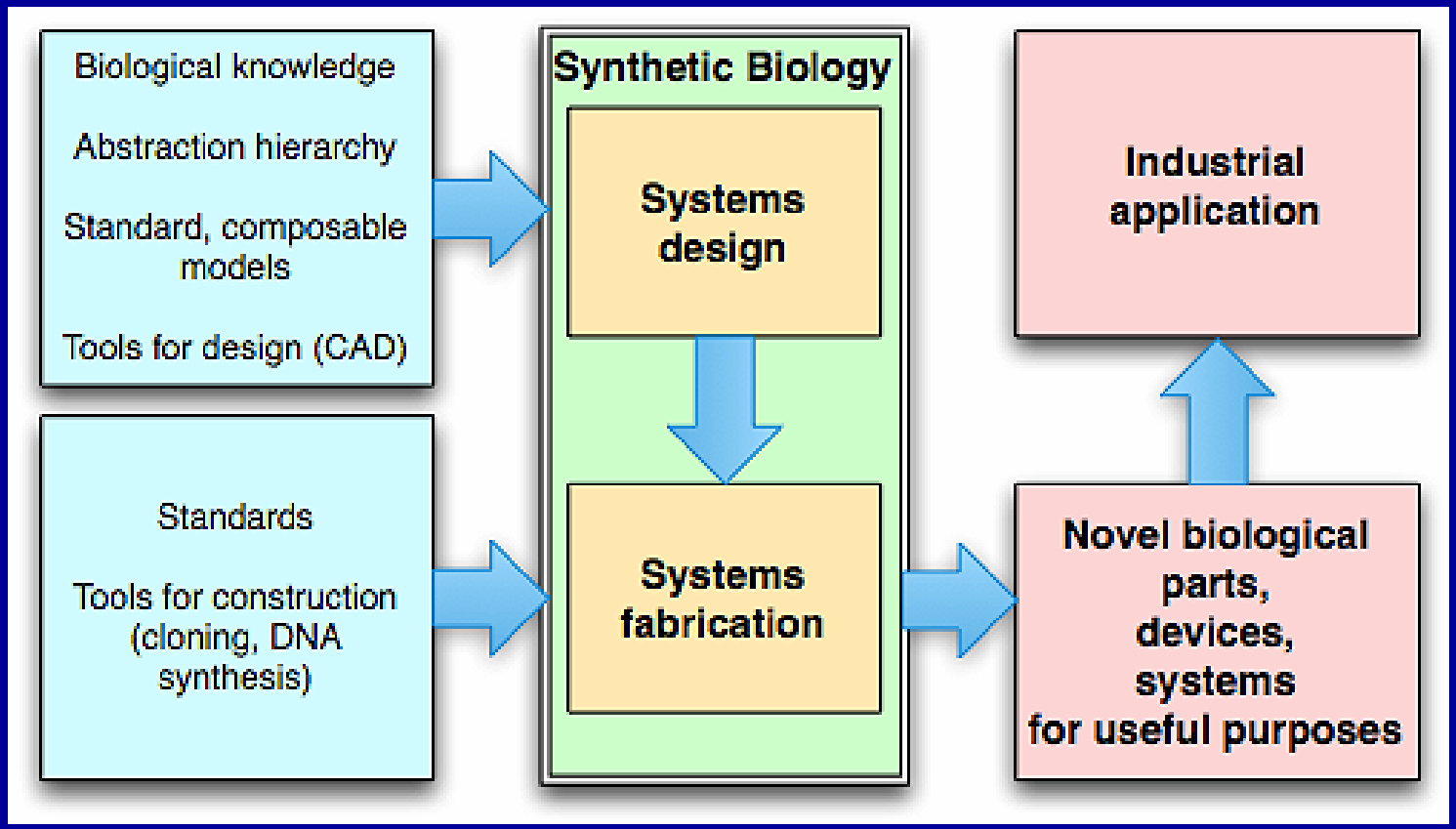
The basis of most engineering - parts, devices and systems

The Engineering Design Cycle

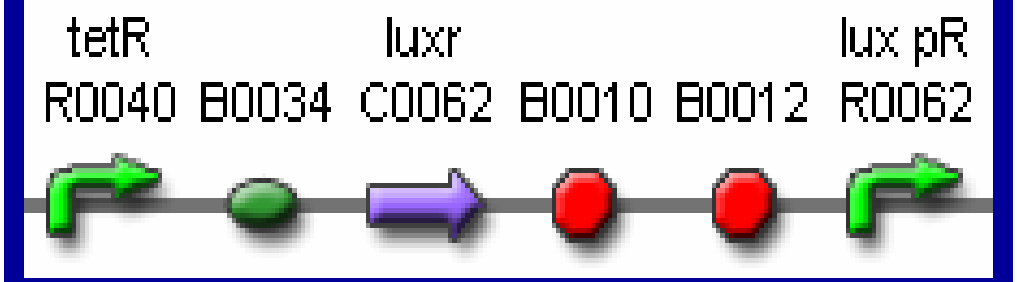


Engineering Biology

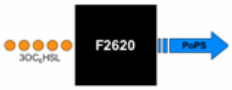
Using the engineering design cycle, the aim is to develop synthetic biology applications following the same efficient workflow that other engineering disciplines successfully use



Biobrick BBa_F2620



BBa_F2620



3OC$_6$HSL → PoPS Receiver

http://parts.mit.edu/registry/index.php/Part:BBa_F2620

Authors:
Barry Canton [bcanton@mit.edu]
Anna Labno [labnoa@mit.edu]

Last Update: 5 October 2006

Description

A transcription factor (LuxR, BBa_C0062) that is active in the presence of cell-cell signaling molecule 3OC$_6$HSL is controlled by a TetR-regulated operator (BBa_R0040). Device input is 3OC$_6$HSL. Device output is PoPS from a LuxR-regulated operator. If used in a cell containing TetR then a second input signal such as aTc can be used to produce a Boolean AND function.

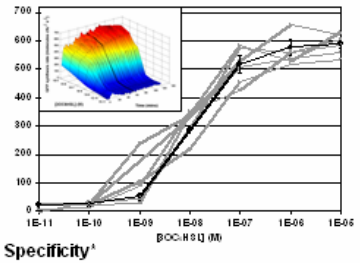
Characteristics

Input Swing: **0.1 to 1000** nM 3OC$_6$HSL, exogenous
Output Swing: **21±3 to 590±9** GFP molecules cfu^{-1} s^{-1}
Switch Point: **10 nM** 3OC$_6$HSL, exogenous
LH Response: **9.7 min** ($t_{50\%}$), **17 min** ($t_{90\%}$)

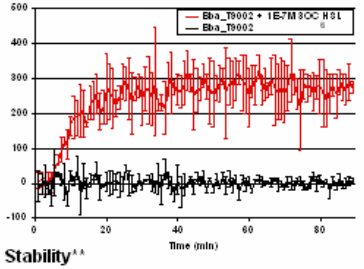
Key Components

BBa_R0040: TetR-regulated operator
BBa_C0062: luxR ORF
BBa_R0062: LuxR-regulated operator

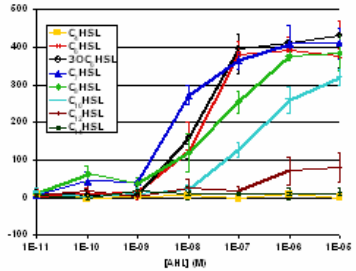
Transfer Function*



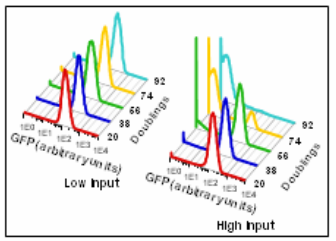
Response Time*



Specificity*



Stability**



Demand (low/high input)

Translational: **336/9449** ribosomes cfu^{-1}
5040/141600 charged tRNA cfu^{-1} s^{-1}

Compatibility

Chassis: Compatible with *MC4100*, *MG1655*, and *DH5α*
Plasmids: Compatible with *pSB3K3* and *pSB1A2*
Devices: Compatible with *E0240*, *E0430* and *E0434*
Crosstalk with systems containing *TetR (C0040)*
Signaling: Crosstalk with input molecules similar to *3OC$_6$HSL*

Stability (low/high input)

Genetic: >**92/74** replication events*
Performance: >**92/74** replication events*

Conditions (abridged)

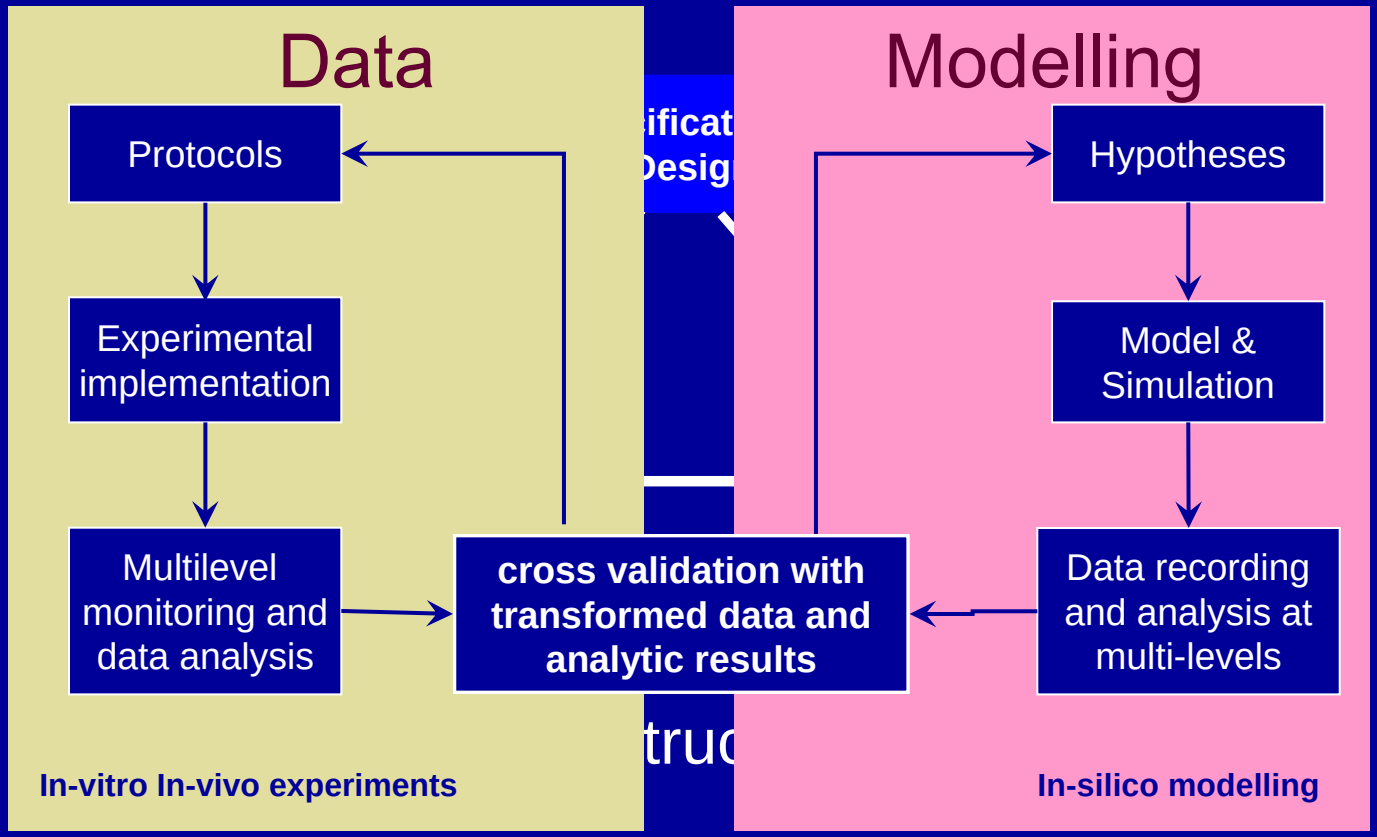
Output: Indirect via *BBa_E0240*
Vector: pSB3K3
Chassis: MG1655
Culture: Supplemented M9, 37°C
**Equipment:* PE Victor3 plate reader
***Equipment:* BD FACScan cytometer

Registry of Standard Biological Parts
making life better, one part at a time

License: Public

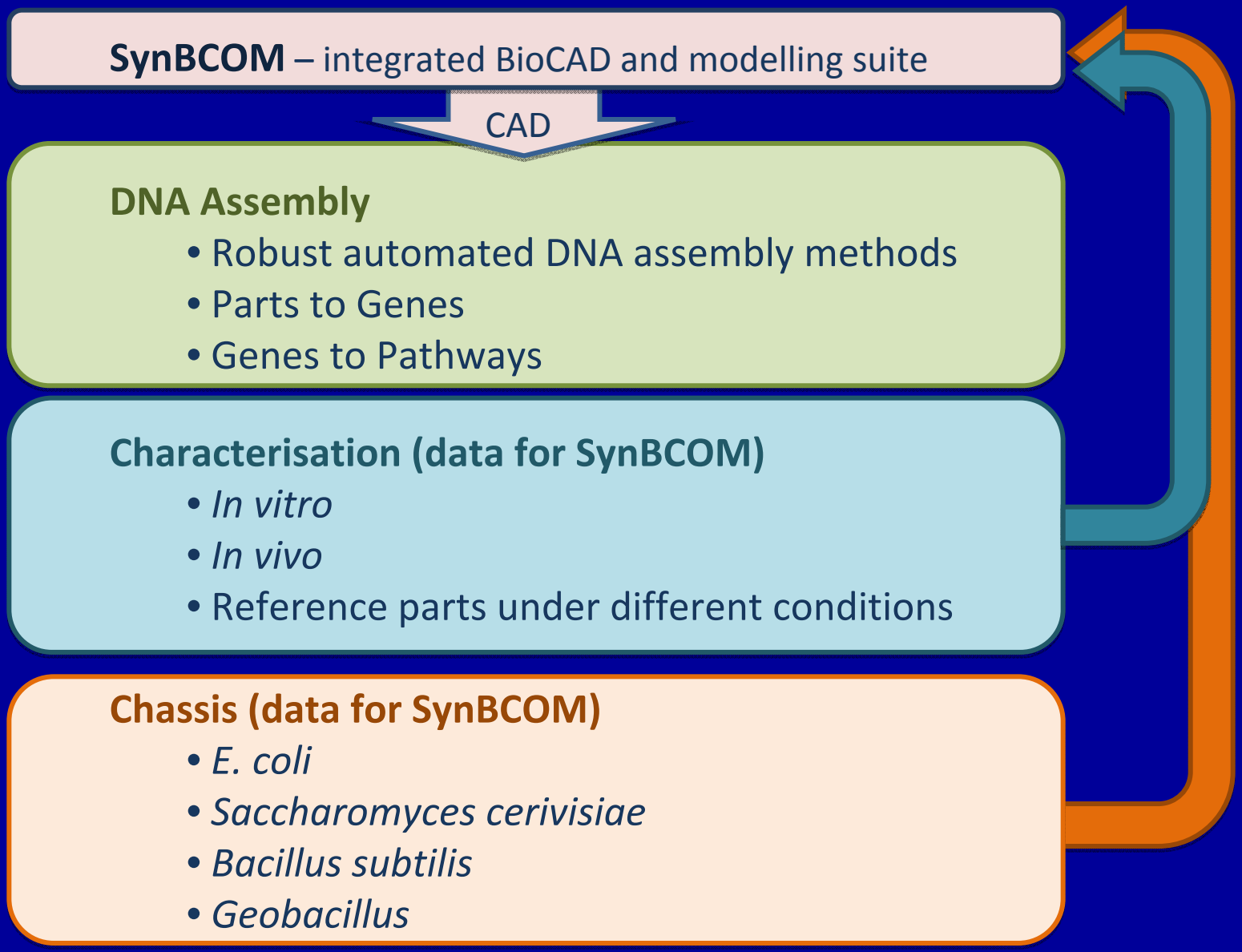
Signaling Devices

A Framework for Synthetic Biology Device/Systems Design

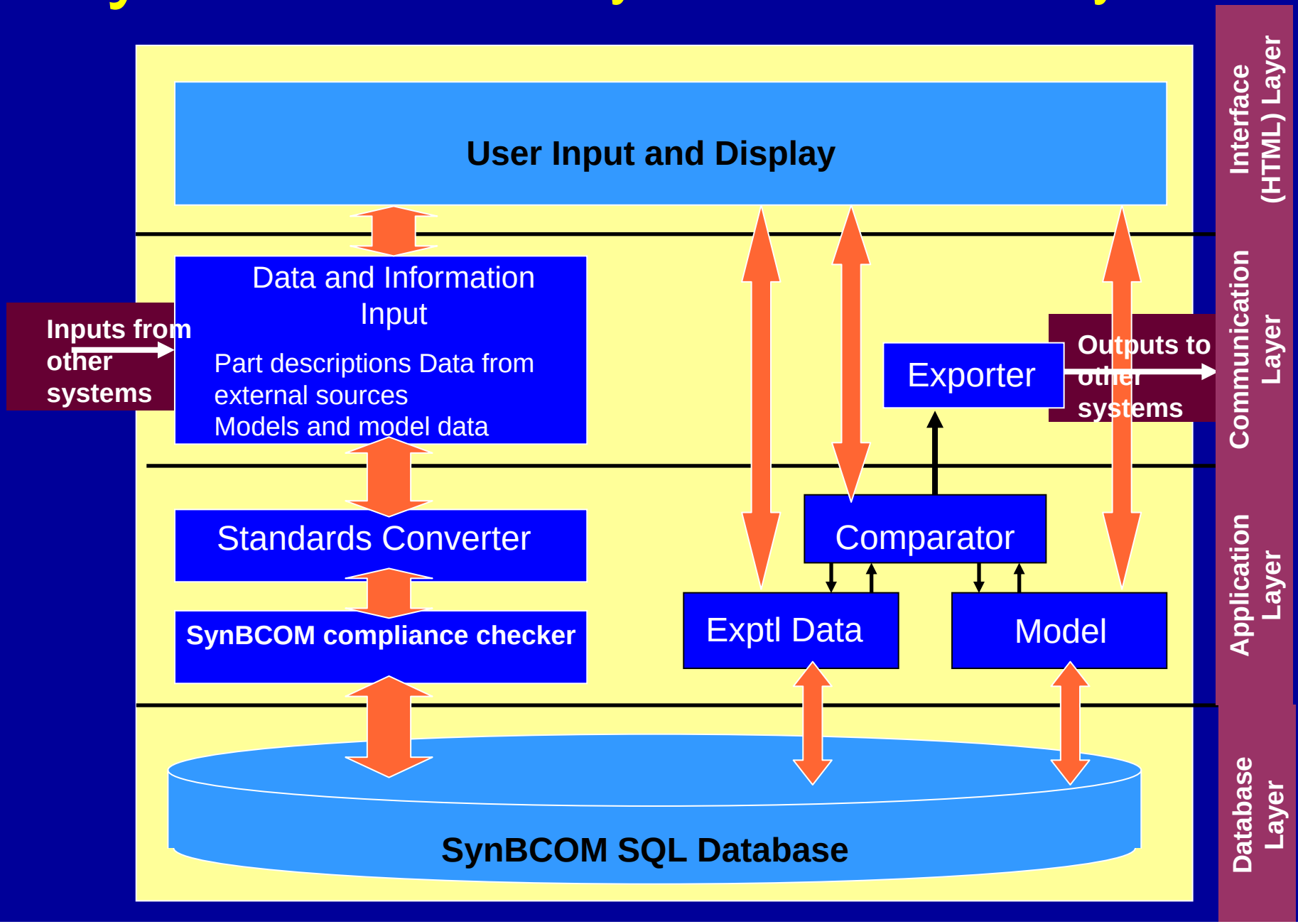


B) Cross validation between model predictions and experimental realities

Foundation Technologies



SynBCOM - a SynB Information System



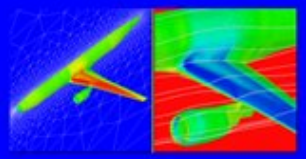
Specification

Part and Device Specification

Design

Part and Device Characterisation, and Design

Modelling



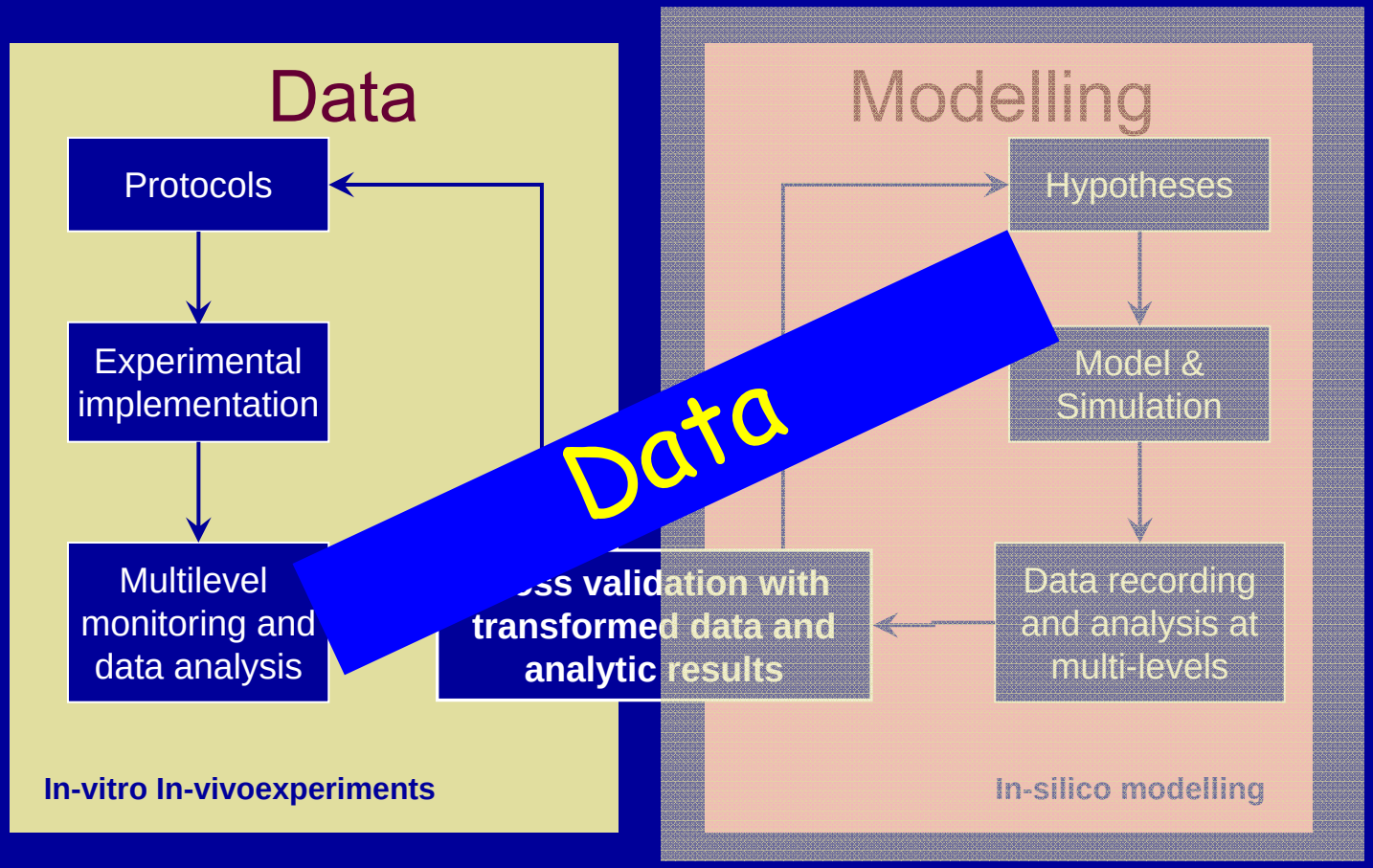
Implementation, Testing and Validation

Small Scale Assembly of Parts and Devices in House

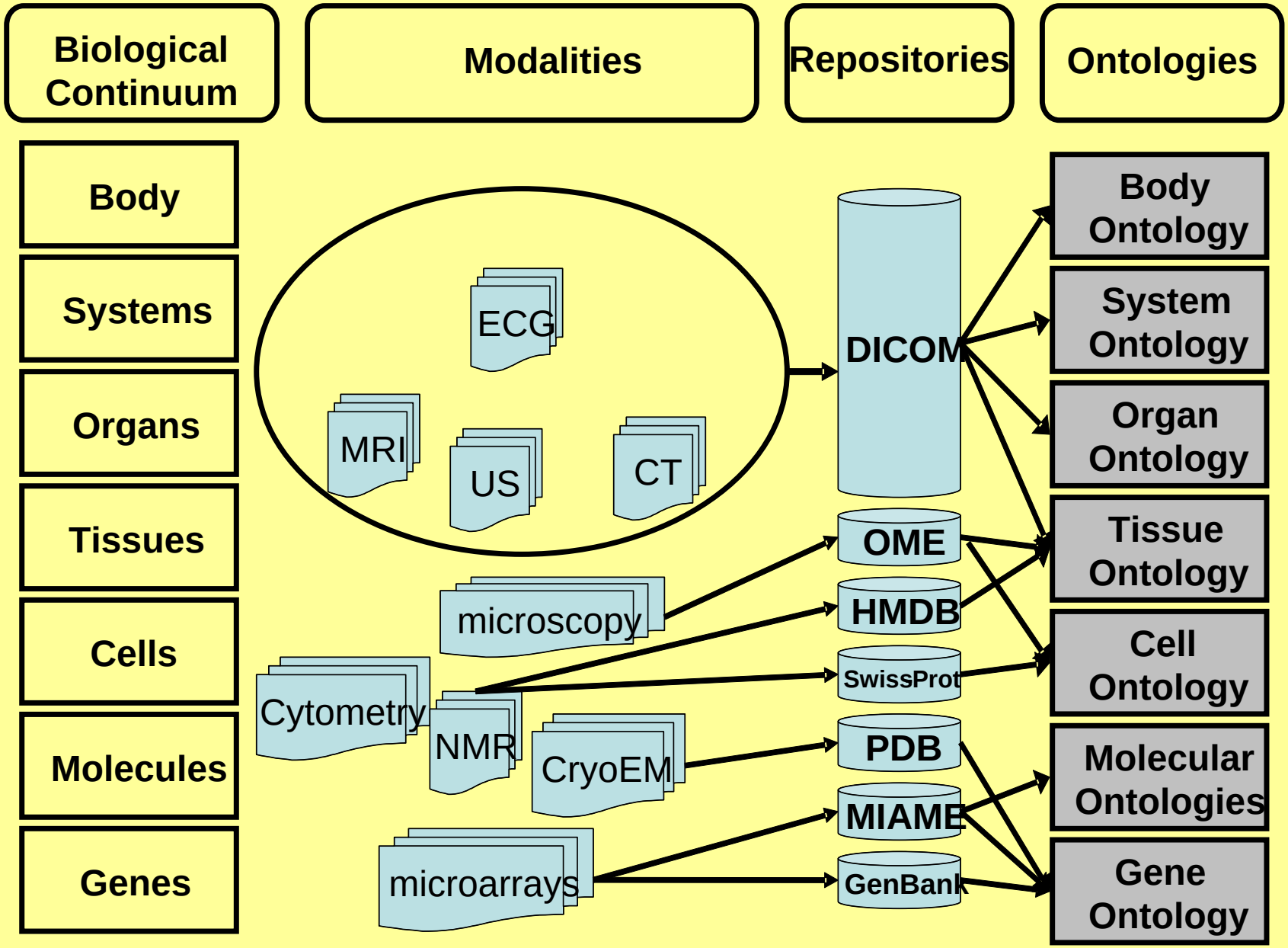
Large Scale Assembly of Parts and Devices within Gene Synthesis Companies

Applications Companies

- Healthcare
- Pharma
- Biofuels
- Agroscience


Data
Protocols
Experimental implementation
Multilevel monitoring and data analysis
In-vitro In-vivoexperiments
Modelling
Hypotheses
Model & Simulation
Data recording and analysis at multi-levels
In-silico modelling
Data
transformed data and
analytic results

Standards


Biological Continuum
Modalities
Repositories
Ontologies
Body
Systems
Organs
Tissues
Cells
Molecules
Genes
ECG
MRI
US
CT
microscopy
Cytometry
NMR
CryoEM
microarrays
DICOM
OME
HMDB
SwissProt
PDB
MIAME
GenBank
Body Ontology
System Ontology
Organ Ontology
Tissue Ontology
Cell Ontology
Molecular Ontologies
Gene Ontology

DICOM

- The standard for Digital Imaging & Communications in Medicine
- Developed by the National Electrical Manufacturers Association (NEMA) & the American College of Radiology (ACR)
- Covers most image formats for all of biomedicine
- Specification for messaging and communication between scanners etc

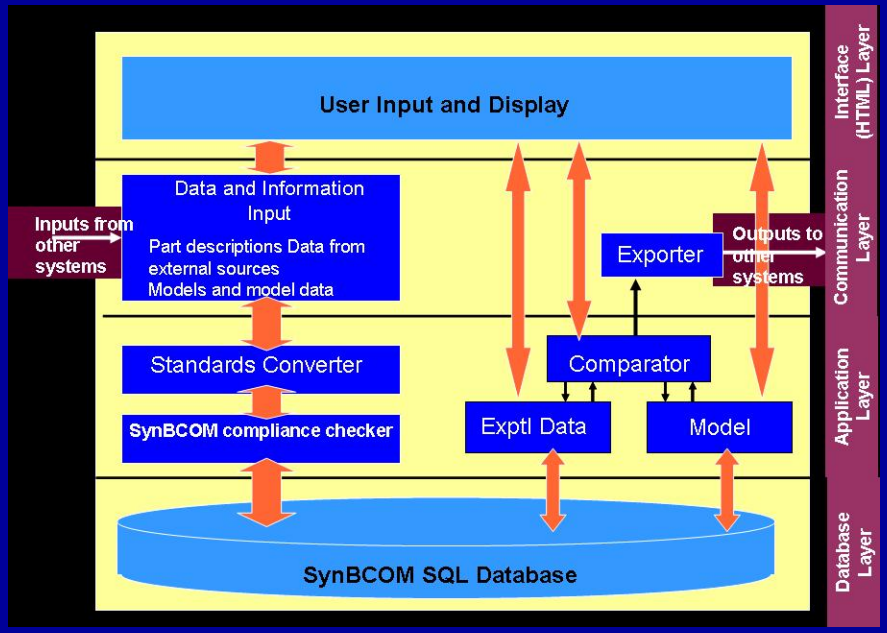
SynBCOM

Based on the DICOM standard for medical images

Machine readable to allow programmes to collate, search and update the information contained where appropriate

Parts will be ontologically organised to aid design

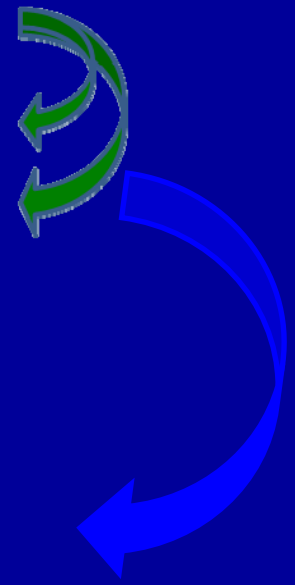
Parts defined by their characteristics, which are determined by experiments and data which will be associated with the part



DICOM in SynB

DICOM can be used & modified to:

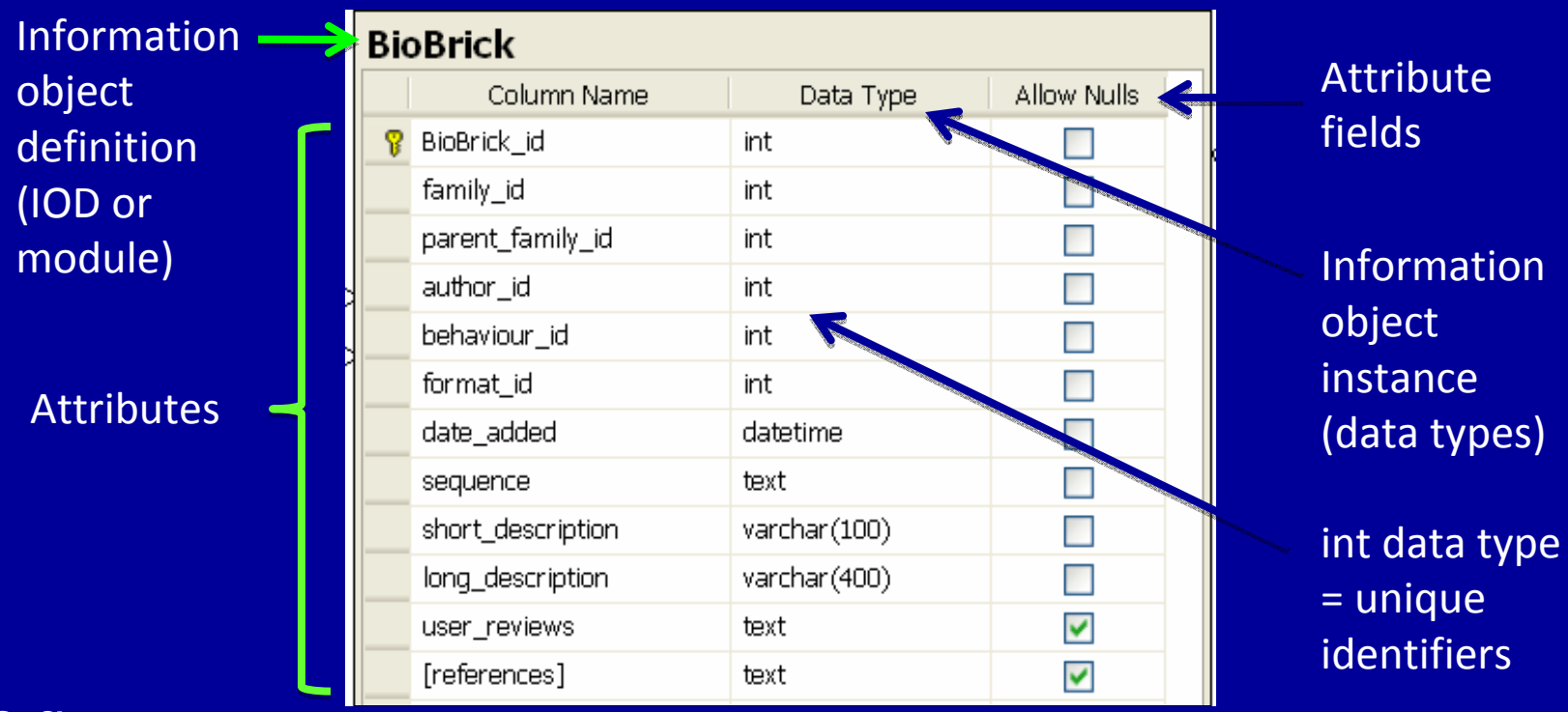
- Store/send parts & associated metadata
- → " related images
- → " related/collected data



- e.g. BioBrick part source, design notes, vector resistance, protocols

- Institution name, date, authors

The DICOM Data Structure

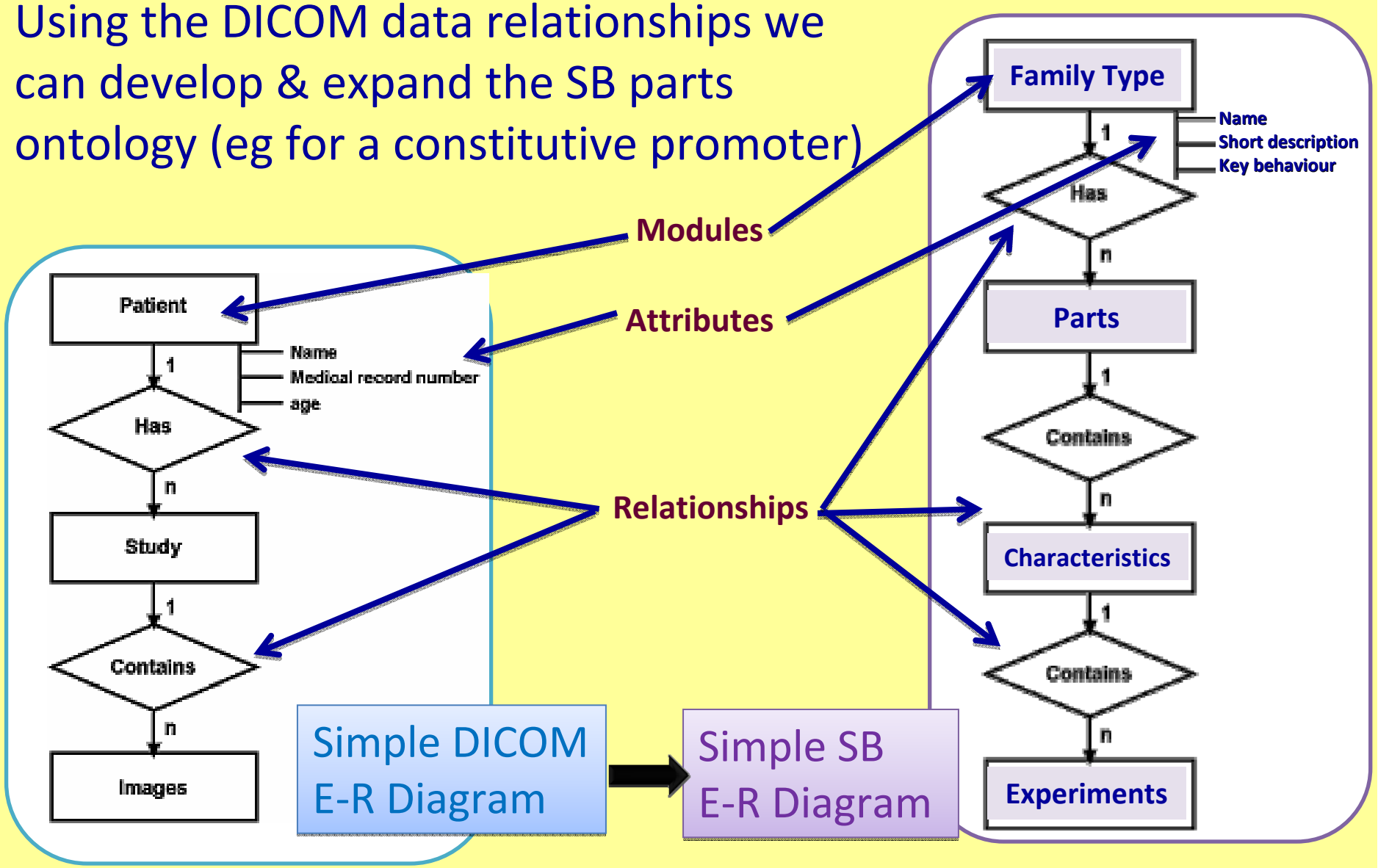


e.g.

- In standard DICOM, each image type (x-ray, ultrasound), will contain some overlapping modules and some specific modules
- Likewise in SB, each family type (constitutive, inducible promoters), will contain some overlapping modules (see table above) and some specific modules

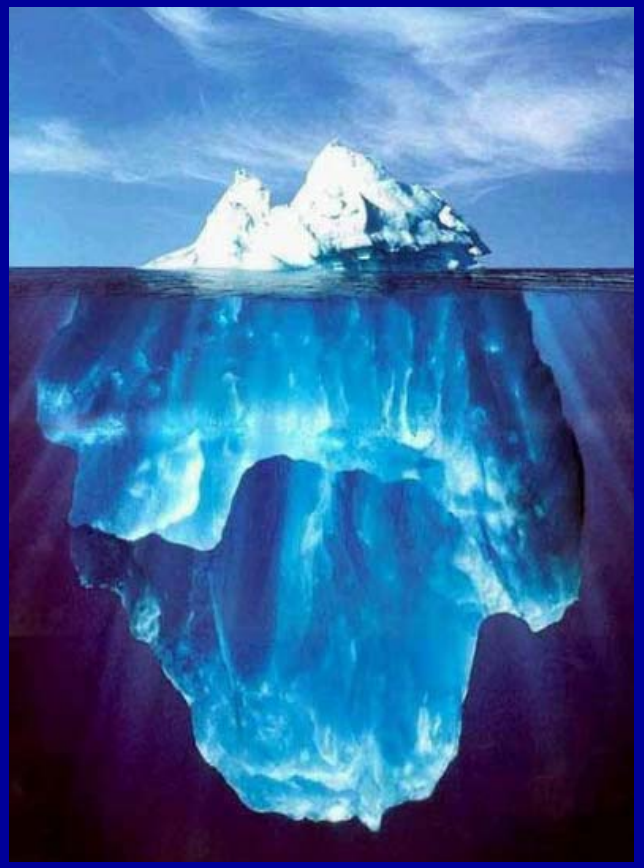
Using DICOM Data Structure

Using the DICOM data relationships we can develop & expand the SB parts ontology (eg for a constitutive promoter)

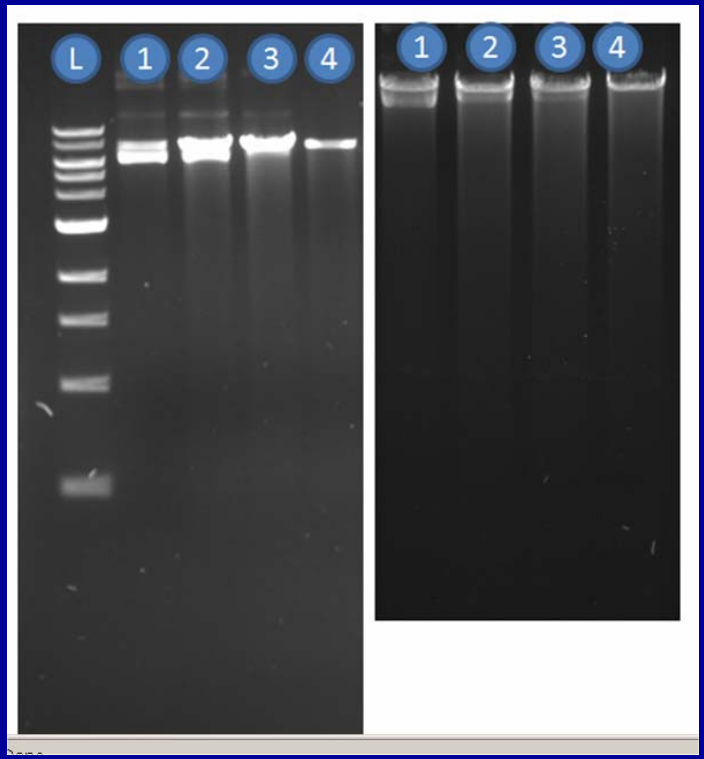


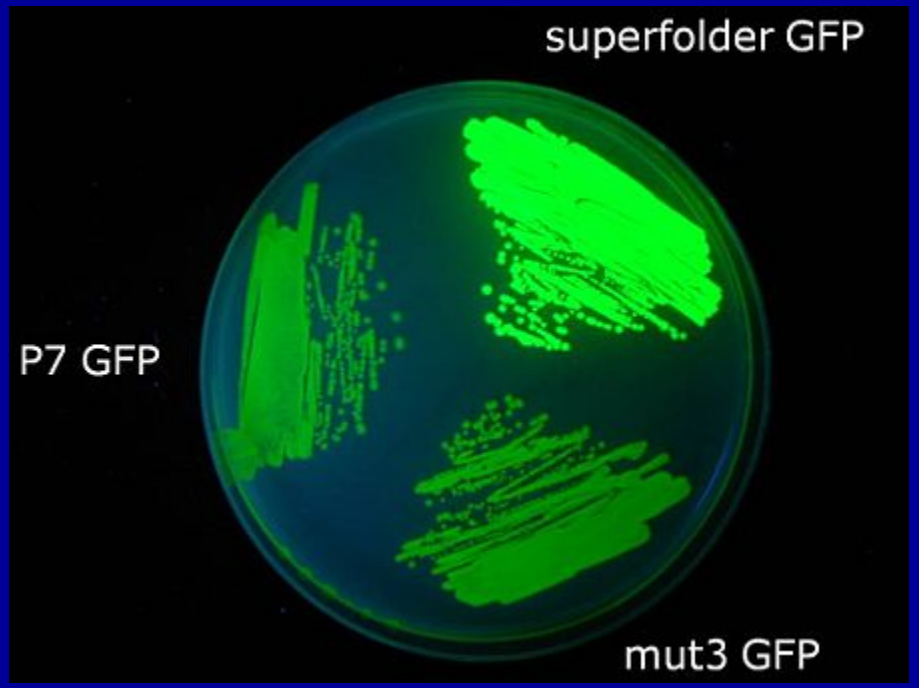
So why DICOM?

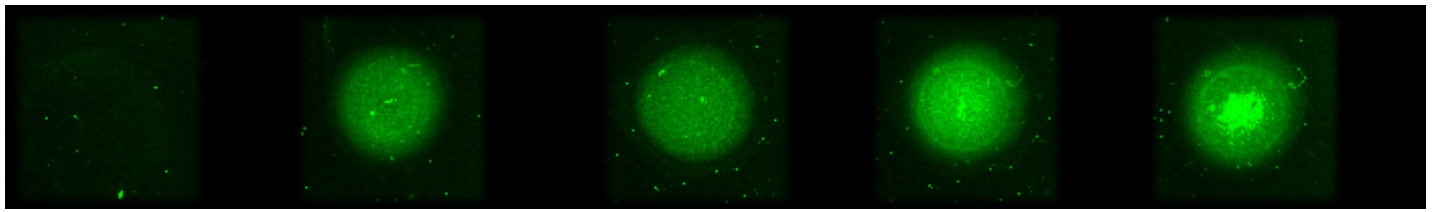
- DICOM is a hugely successful international standard that in the biomedical field
- Entirely novel to SB
- Each image is attached to its own metadata
- A lot of time and effort has gone into developing DICOM
- Modular structure
- Images (colour accuracy)
- Modality friendly
- Includes communications standards

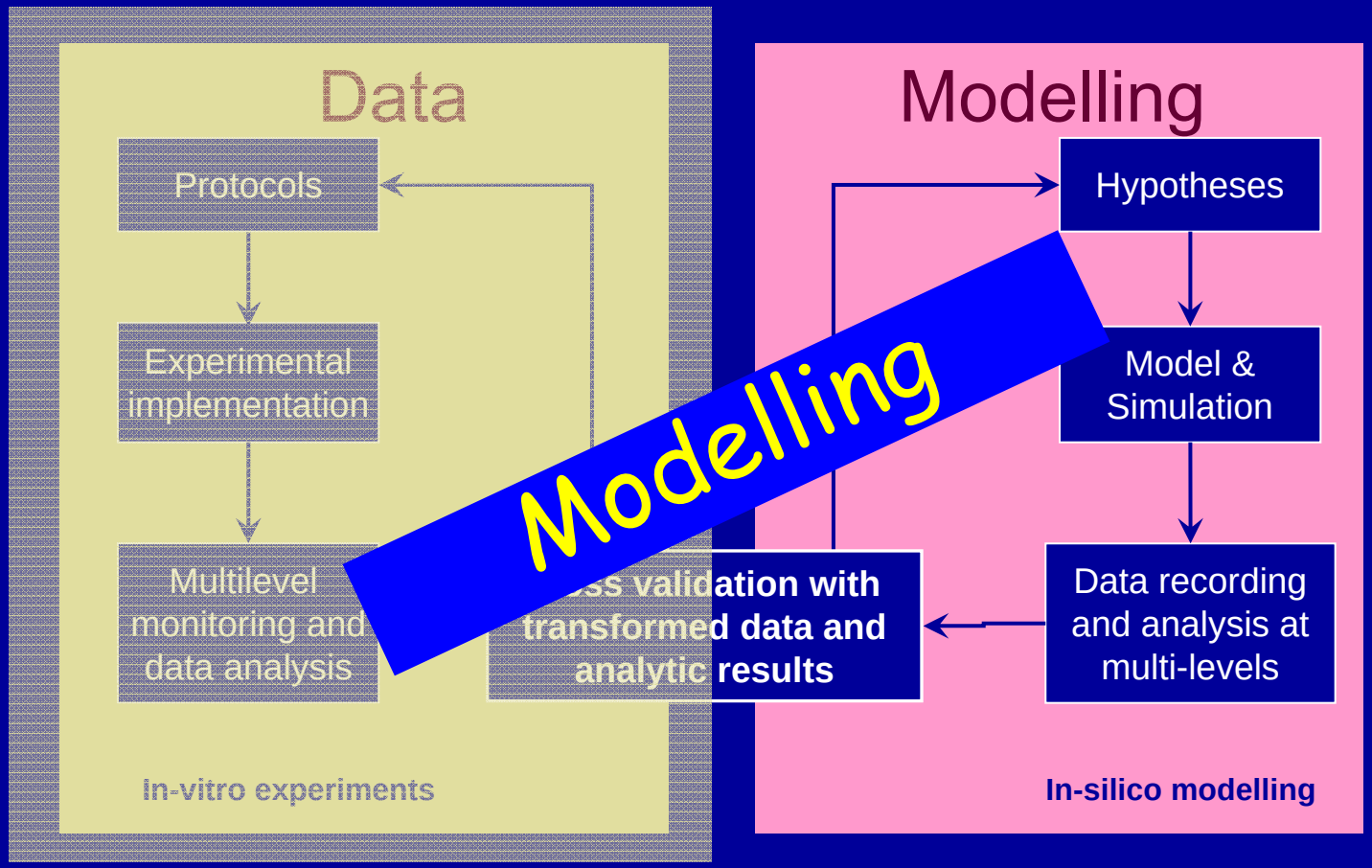


One Example – Colour Accuracy








Data
Protocols
Experimental implementation
Multilevel monitoring and data analysis
In-vitro experiments
Modelling
Hypotheses
Model & Simulation
Data recording and analysis at multi-levels
In-silico modelling
Modelling

Developing a Registry of standard, composable models

Current tools

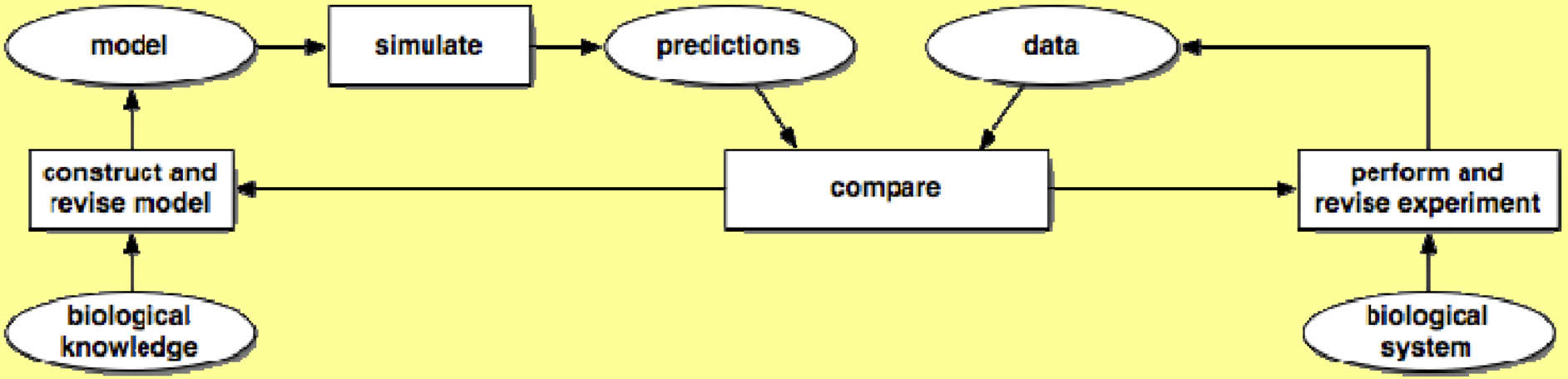
- There are already many systems biology model repositories (e.g., Biomodels, CellML model repository, Open Wetware repository, Java web simulation online, ModelDB, etc.) and *model analysis and design tools* available

- However, these repositories and tools lack some of the important features of a proper SynB CAD framework

- They hardly support the modular building process used to create complex systems from the interconnection of parts and forming an integral part of the engineering cycle

- They do not provide a unified CAD environment with access to composable and reusable mathematical models

Model Based Analysis



This is the classical approach used in *systems biology*

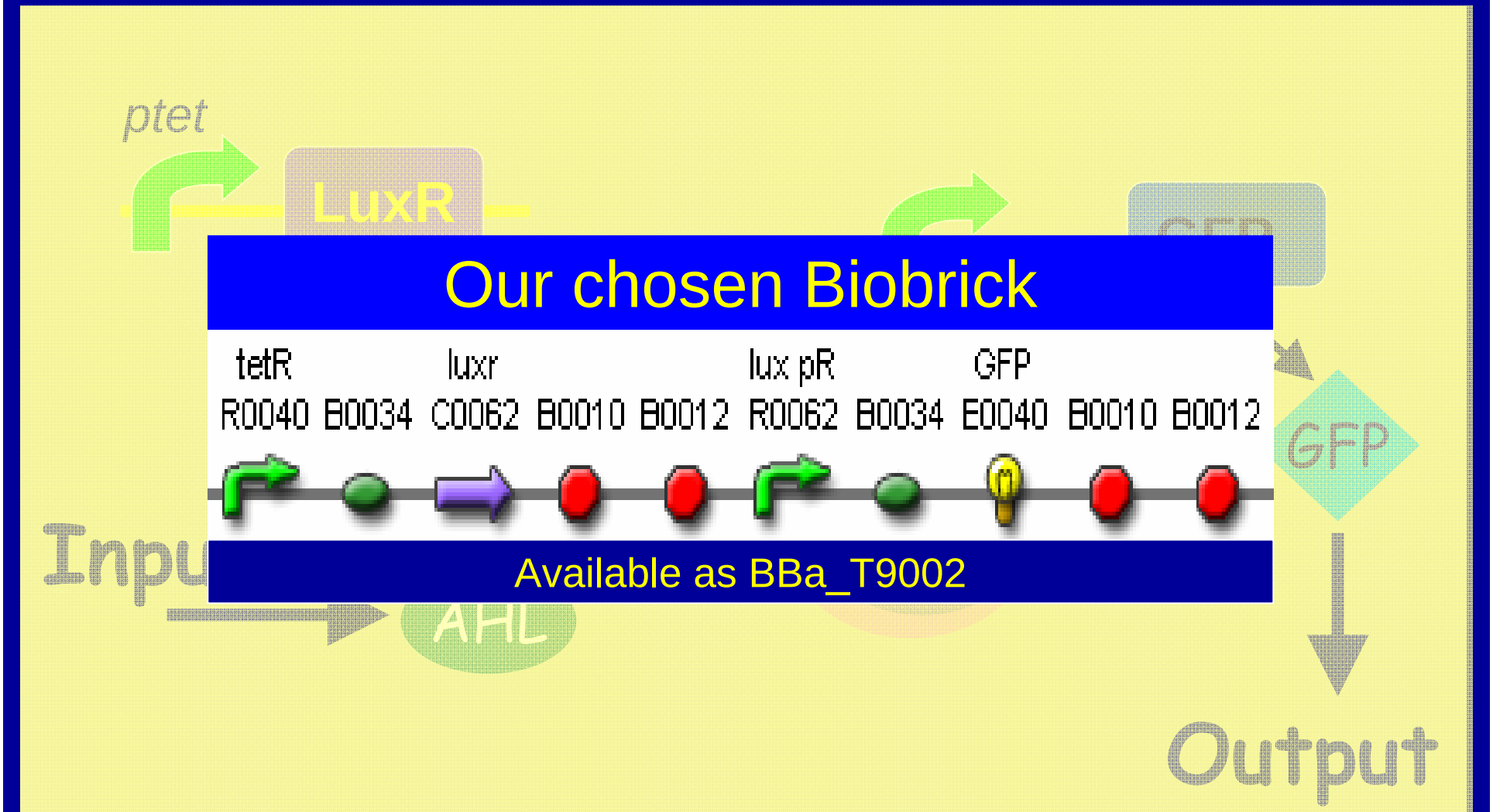
It is appropriate for the development of *monolithic* models

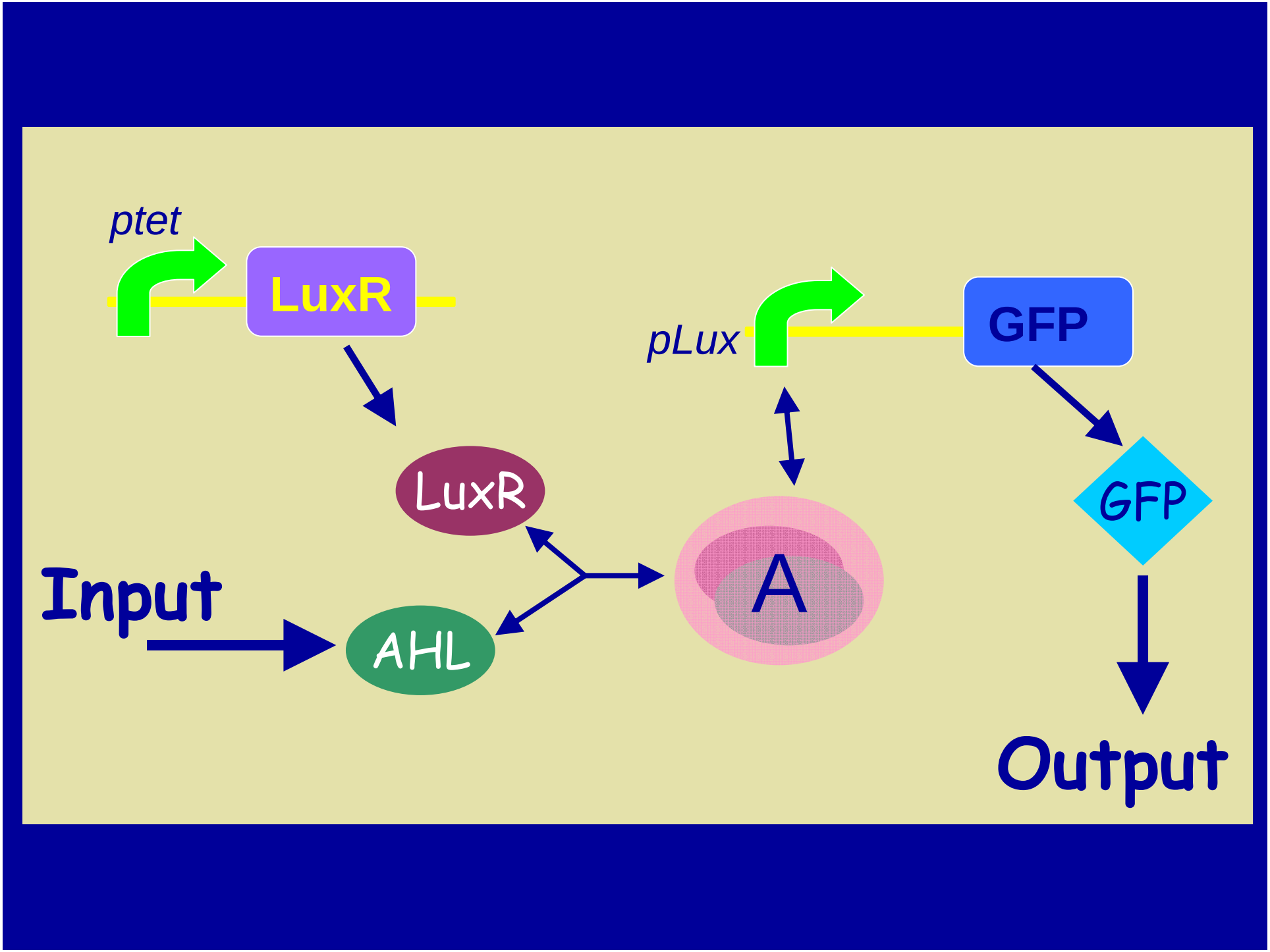
The development of CAD tools for synthetic biology requires more (modularity, composability, database integration)

Infector Detector

iGEM 2007

The Biochemical Network – the basis of Infector Detector




ptet
LuxR
pLux
GFP
LuxR
GFP
Input
A
AHL
Output

Modelling

$$\frac{d[LuxR]}{dt} = K_1\mu(R) + K_3[A] - K_2[LuxR][AHL]$$

$$\frac{d[AHL]}{dt} = K_3[A] - K_2[LuxR][AHL]$$

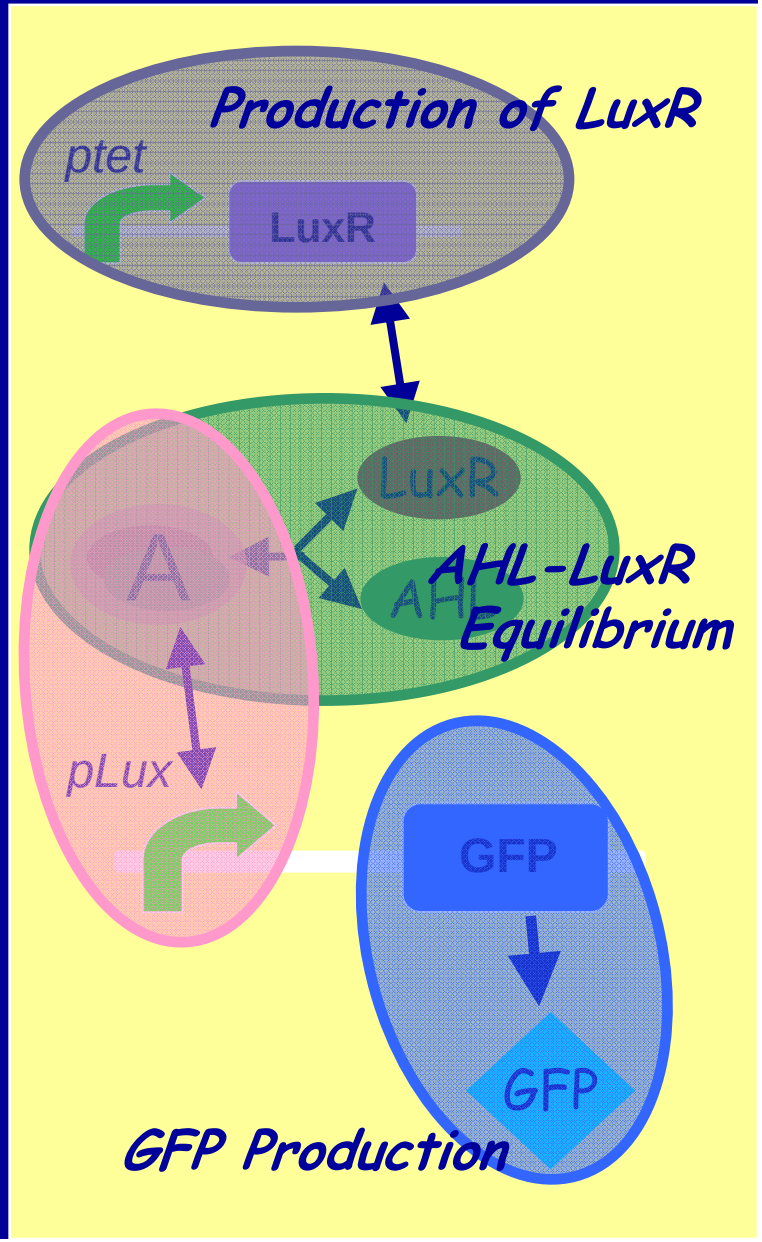
$$\frac{d[A]}{dt} = -K_3[A] + K_2[LuxR][AHL] + K_5[AP] - K_4[A][P]$$

$$\frac{d[P]}{dt} = -\frac{d[AP]}{dt} = K_5[AP] - K_4[A][P]$$

$$\frac{d[GFP]}{dt} = K_6[AP]\mu(R)$$

$$\frac{d[R]}{dt} = -\alpha_1 K_1\mu(R) - \alpha_2 K_6[AP]\mu(R)$$

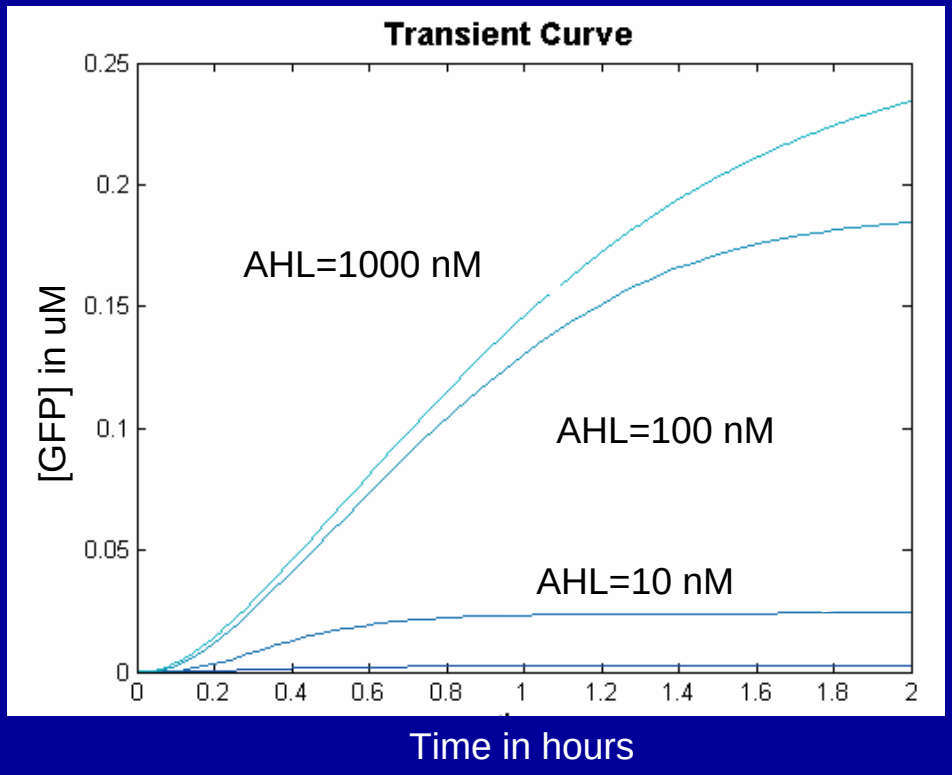
$$\text{Curbing Function } \mu(R) = \frac{R^p}{1+R^p} \text{ (Hill Function)}$$



Typical Simulations

General Behaviour:

- Slow uptake
- Saturation after few hours (Resources exhausted)
- The higher the input (AHL) , the higher the output (GFP)

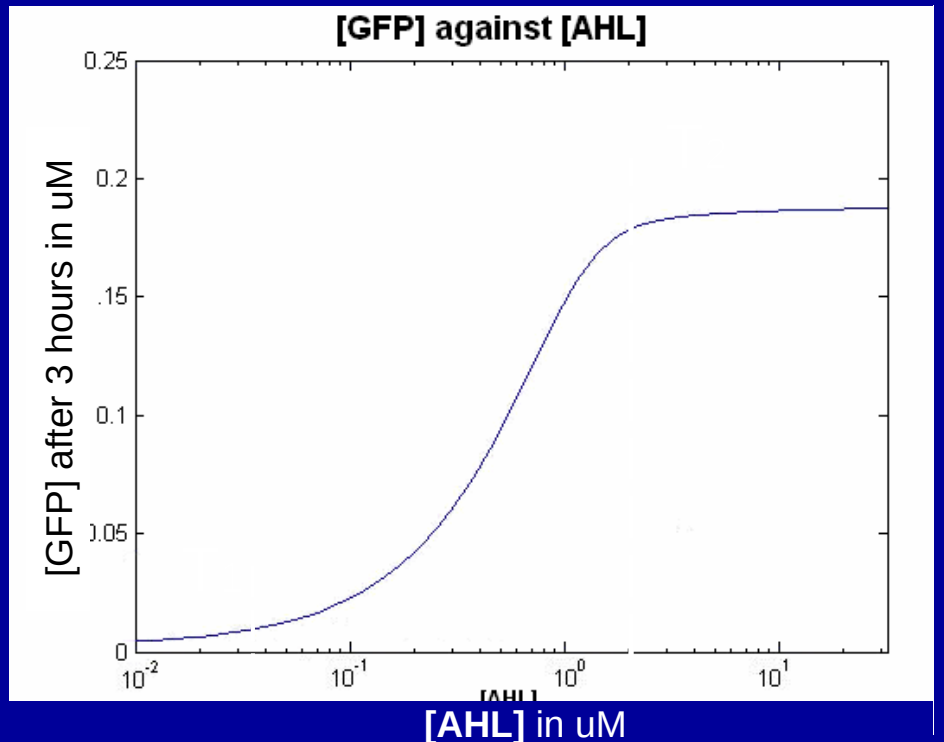


Transfer Function

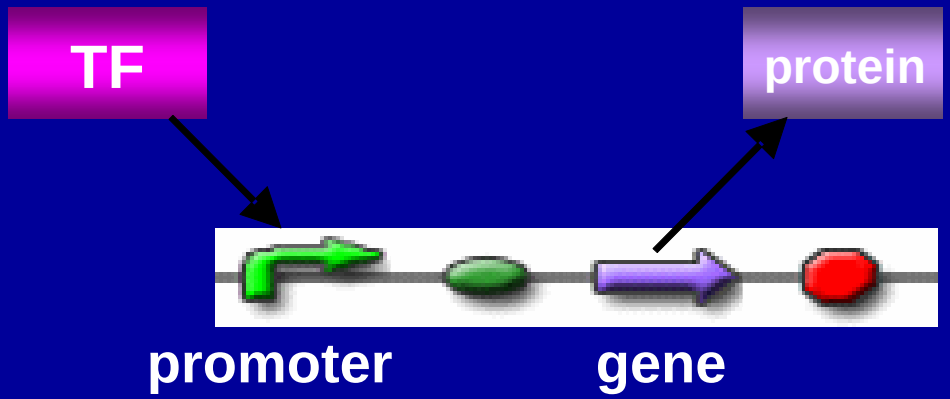
AHL → **Infector Detector** → **GFP**

GFP vs AHL

- Similar to F2620 in vivo
- Below T_1 : No detection
- Above T_2: Saturation



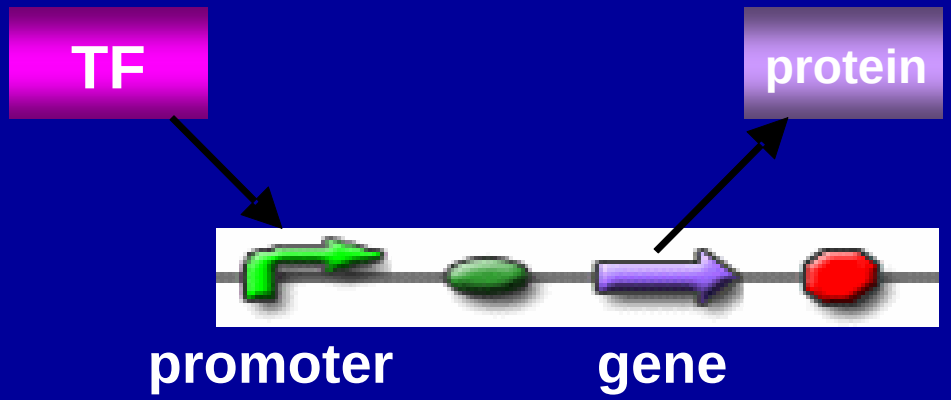
A typical transcriptional regulatory device



$$\frac{d[mRNA]}{dt} = \frac{k_{tr} \cdot \left(\frac{W^n}{K^n}\right)^{\mu}}{1 + \left(\frac{W^n}{K^n}\right)} - d_m.[mRNA]$$

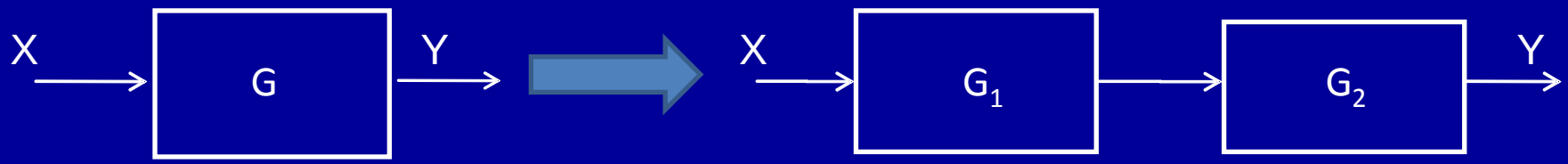
Currently ODEs are mainly used for modelling in Synthetic Biology

This becomes cumbersome as the complexity of the systems increases

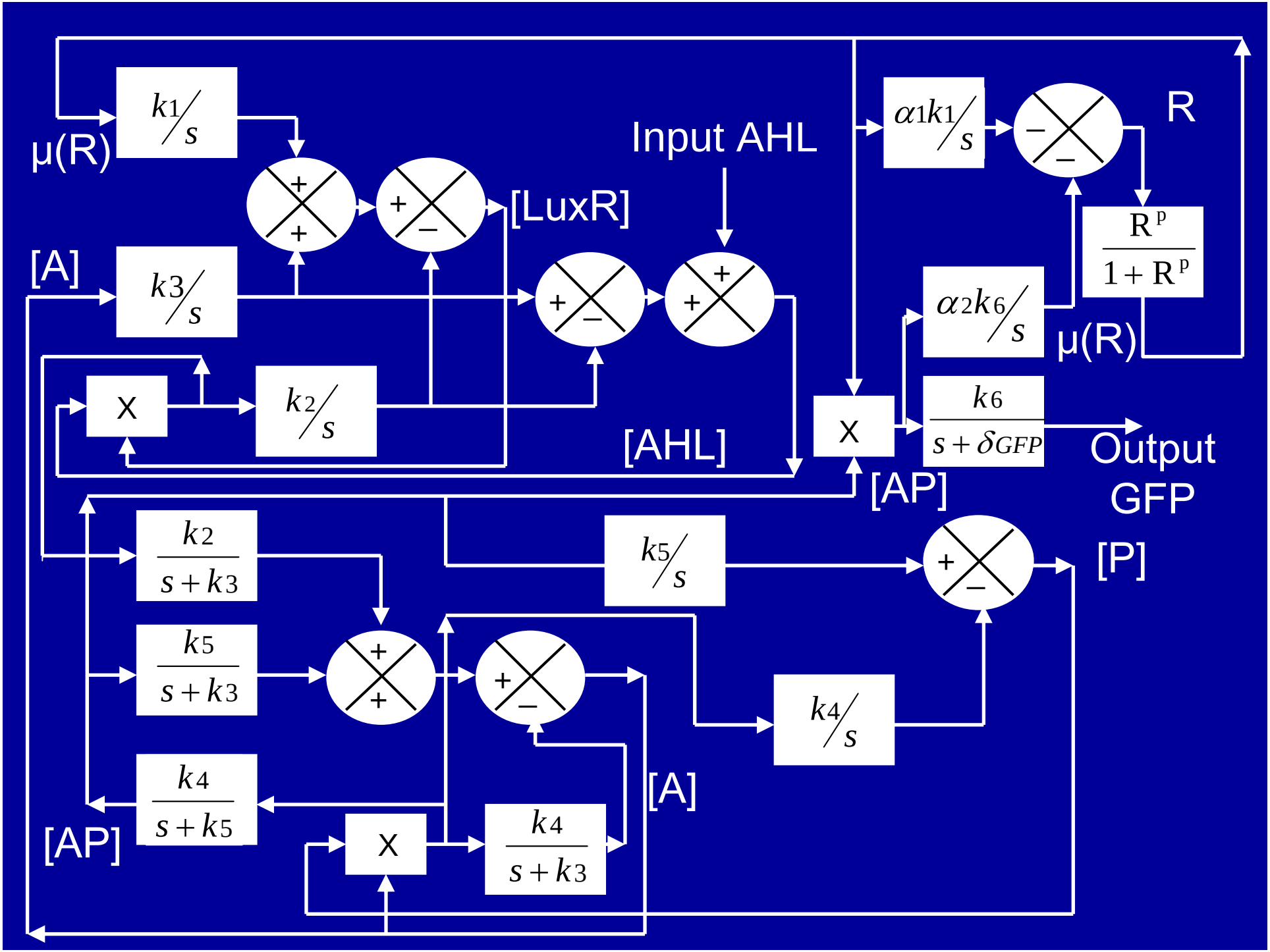


What is required is the application of Systems Theory

Modularisation



and, the application of Transform Methods


μ(R)
k_1/s
[A]
k_3/s
X
k_2/s
+
+
+
−
[LuxR]
Input AHL
+
+
+
−
[AHL]
$\alpha_1 k_1/s$
−
−
R
$\frac{R^p}{1+R^p}$
$\alpha_2 k_6/s$
μ(R)
X
[AP]
$\frac{k_6}{s+\delta_{GFP}}$
Output
GFP
$\frac{k_2}{s+k_3}$
$\frac{k_5}{s+k_3}$
$\frac{k_4}{s+k_5}$
[AP]
+
+
+
−
X
$\frac{k_4}{s+k_3}$
k_5/s
+
−
[P]
k_4/s
[A]

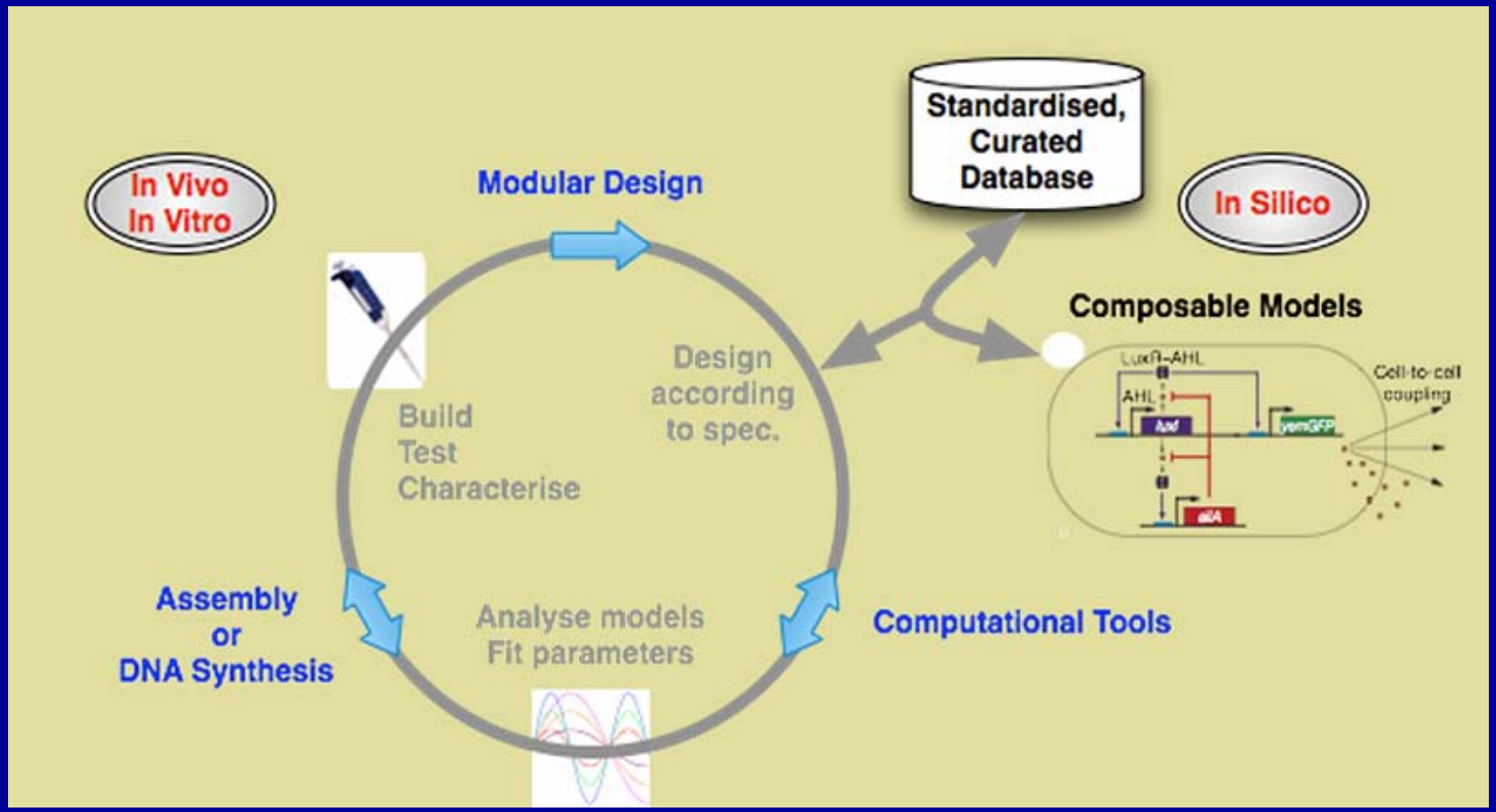
So, what is needed ?

A modular In Silico CAD framework allowing:

- Easy design, simulation and composition of SynB models
- Direct robustness and sensitivity analysis of models
- Seamless integration with a standardised & curated database:
  - search & annotation of part models based on design specification
  - search & modulation of model parameters
  - automated DNA sequence prediction & *de novo* synthesis

CAD and a Professional Model Registry

The Engineering Design Cycle



The End